Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2133

An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators[1]

By Cornelius Lanczos

The present investigation designs a systematic method for finding the latent roots and the principal axes of a matrix, without reducing the order of the matrix. It is characterized by a wide field of applicability and great accuracy, since the accumulation of rounding errors is avoided, through the process of "minimized iterations". Moreover, the method leads to a well convergent successive approximation procedure by which the solution of integral equations of the Fredholm type and the solution of the eigenvalue problem of linear differential and integral operators may be accomplished.

I. Introduction

The eigenvalue problem of linear operators is of central importance for all vibration problems of physics and engineering. The vibrations of elastic structures, the flutter problems of aerodynamics, the stability problem of electric networks, the atomic and molecular vibrations of particle physics, are all diverse aspects of the same fundamental problem, viz., the principal axis problem of quadratic forms.

In view of the central importance of the eigenvalue problem for so many fields of pure and applied mathematics, much thought has been devoted to the designing of efficient methods by which the eigenvalues of a given linear operator may be found. That linear operator may be of the algebraic or of the continuous type; that is, a matrix, a differential operator, or a Fredholm kernel function. Iteration methods play a prominent part in these designs, and the literature on the iteration of matrices is very extensive.[2] In the English literature of recent years the works of H. Hotelling [1][3] and A. C. Aitken [2] deserve attention. H. Wayland [3] surveys the field in its historical development, up to recent years. W. U. Kincaid [4] obtained additional results by improving the convergence of some of the classical procedures.

The present investigation, although starting out along classical lines, proceeds nevertheless in a different direction. The advantages of the method here developed[4] can be summarized as follows:

1. The iterations are used in the most economical fashion, obtaining an arbitrary number of eigenvalues and eigensolutions by one single set of iterations, without reducing the order of the matrix.

2. The rapid accumulation of fatal rounding errors, common to all iteration processes if applied to matrices of high dispersion (large "spread" of the eigenvalues), is effectively counteracted by the method of "minimized iterations".

3. The method is directly translatable into analytical terms, by replacing summation by integration. We then get a rapidly convergent analytical iteration process by which the eigenvalues and eigensolutions of linear differential and integral equations may be obtained.

II. The Two Classical Solutions of Fredholm's Problem

Since Fredholm's fundamental essay on integral equations [5], we can replace the solution of linear differential and integral equations by the solution

[1] The preparation of this paper was sponsored in part by the Office of Naval Research.

[2] The basic principles of the various iteration methods are exhaustively treated in the well-known book on Elementary matrices by R. A. Frazer, W. J. Duncan, and A. R. Collar (Cambridge University Press, 1938); (MacMillan, New York, N. Y., 1947).

[3] Figures in brackets indicate the literature references at the end of this paper.

[4] The literature available to the author showed no evidence that the methods and results of the present investigation have been found before. However, A. M. Ostrowski of the University of Basle and the Institute for Numerical Analysis informed the author that his method parallels the earlier work of some Russian scientists; the references given by Ostrowski are: A. Krylov, Izv. Akad. Nauk SSSR 7, 491 to 539 (1931); N. Luzin, Izv. Akad. Nauk SSSR 7, 903 to 958 (1931). On the basis of the reviews of these papers in the Zentralblatt, the author believes that the two methods coincide only in the point of departure. The author has not, however, read these Russian papers.

904428—50——1

UNIVERSITY OF MICHIGAN LIBRARIES

of a set of simultaneous ordinary linear equations of infinite order. The problem of Fredholm, if formulated in the language of matrices, can be stated as follows: Find a solution of the equation

$$y-\lambda Ay=b, \tag{1}$$

where b is a given vector, λ a given scalar parameter, and A a given matrix (whose order eventually approaches infinity); whereas y is the unknown vector. The problem includes the inversion of a matrix ($\lambda=\infty$) and the problem of the characteristic solutions, also called "eigensolutions", ($b=0$) as special cases.

Two fundamentally different classical solutions of this problem are known. The first solution is known as the "Liouville-Neumann expansion" [6]. We consider A as an algebraic operator and obtain formally the following infinite geometric series:

$$y=\frac{1}{1-\lambda A}\,b=(1+\lambda A+\lambda^2A^2+\ \ldots)b. \tag{2}$$

This series converges for sufficiently small values of $|\lambda|$ but diverges beyond a certain $|\lambda|=|\lambda_1|$. The solution is obtained by a series of successive "iterations";[5] we construct in succession the following set of vectors:

$$\left.\begin{aligned} b_0&=b\\ b_1&=Ab_0\\ b_2&=Ab_1\\ &\cdot\\ &\cdot\\ &\cdot\\ b_{n+1}&=Ab_n \end{aligned}\right\} \tag{3}$$

and then form the sum:

$$y=b_0+\lambda b_1+\lambda^2b_2+\ \ldots\ . \tag{4}$$

The merit of this solution is that it requires nothing but a sequence of iterations. The drawback of the solution is that its convergence is limited to sufficiently small values of λ.

The second classical solution is known as the Schmidt series [7]. We assume that the matrix A is "nondefective" (i. e. that all its elementary divisors are linear). We furthermore assume that we possess all the eigenvalues[6] μ_i and eigenvectors u_i of the matrix A, defined by the equations

$$Au_i=\mu_iu_i \qquad (i=1,2,\ \ldots\ ,n). \tag{5}$$

If A is nonsymmetric, we need also the "adjoint" eigenvectors u_i^*, defined with the help of the transposed matrix A^*:

$$A^*u_i^*=\mu_iu_i^* \qquad (i=1,2,\ \ldots\ ,n). \tag{6}$$

We now form the scalars

$$\gamma_i=\frac{b\cdot u_i^*}{u_i\cdot u_i^*}, \tag{7}$$

and obtain y in form of the following expansion:

$$y=\frac{\gamma_1u_1}{1-\lambda\mu_1}+\frac{\gamma_2u_2}{1-\lambda\mu_2}+\ \cdots\ +\frac{\gamma_nu_n}{1-\lambda\mu_n}. \tag{8}$$

This series offers no convergence difficulties, since it is a finite expansion in the case of matrices of finite order and yields a convergent expansion in the case of the infinite matrices associated with the kernels of linear differential and integral operators.

The drawback of this solution is—apart from the exclusion of defective matrices[7]—that it presupposes the complete solution of the eigenvalue problem associated with the matrix A.

III. Solution of the Fredholm Problem by the S-Expansion

We now develop a new expansion that solves the Fredholm problem in similar terms as the Liouville-Neumann series but avoids the convergence difficulty of that solution.

We first notice that the iterated vectors b_0, b_1, b_2, ... cannot be linearly independent of each other beyond a certain definite b_k. All these vectors find their place within the n-dimensional space of the matrix A, hence not more than n of them can be linearly independent. We thus know in advance that a linear identity of the following form must exist between the successive iterations.

$$b_m+g_1b_{m-1}+g_2b_{m-2}+\cdots+g_mb_0=0. \tag{9}$$

[5] Throughout this paper the term "iteration" refers to the application of the given matrix A to a given vector b, by forming the product Ab.

[6] We shall use the term "eigenvalue" for the numbers μ_i defined by (5), whereas the reciprocals of the eigenvalues: $\lambda_i=1/\mu_i$ shall be called "characteristic numbers".

[7] The characteristic solutions of defective matrices (i. e. matrices whose elementary divisors are not throughout linear) do not include the entire n-dimensional space, since such matrices possess less than n independent principal axes.

We cannot tell in advance what m will be, except for the lower and upper bounds:

$$1 \leqq m \leqq n. \tag{10}$$

How to establish the relation (9) by a systematic algorithm will be shown in section VI. For the time being we assume that the relation (9) is already established. We now define the polynomial

$$G(x) = x^m + g_1 x^{m-1} + \cdots + g_m, \tag{11}$$

together with the "inverted polynomial" (the coefficients of which follow the opposite sequence):

$$S_m(\lambda) = 1 + g_1\lambda + g_2\lambda^2 + \cdots + g_m\lambda^m. \tag{12}$$

Furthermore, we introduce the partial sums of the latter polynomial:

$$\left.\begin{aligned} S_0 &= 1 \\ S_1(\lambda) &= 1 + g_1\lambda \\ S_2(\lambda) &= 1 + g_1\lambda + g_2\lambda^2 \\ &\cdot \\ &\cdot \\ &\cdot \\ S_{m-1}(\lambda) &= 1 + g_1\lambda + \cdots + g_{m-1}\lambda^{m-1} \end{aligned}\right\}. \tag{13}$$

We now refer to a formula which can be proved by straightforward algebra:[5]

$$\frac{S_m(\lambda) - \lambda^m G(x)}{1 - \lambda x} = S_{m-1}(\lambda) + S_{m-2}(\lambda) \cdot \lambda x + S_{m-3}(\lambda) \cdot \lambda^2 x^2 + \cdots + S_0 \cdot \lambda^{m-1} x^{m-1}. \tag{14}$$

Let us apply this formula operationally, replacing x by the matrix A, and operating on the vector b_0. In view of the definition of the vectors b_i, the relation (9) gives:

$$G(A) \cdot b_0 = 0, \tag{15}$$

and thus we obtain:

$$\frac{S_m(\lambda)}{1 - \lambda A} b_0 = S_{m-1}(\lambda) b_0 + S_{m-2}(\lambda)\lambda b_1 + \cdots + S_0 \lambda^{m-1} b_{m-1}, \tag{16}$$

and hence:

$$y = \frac{1}{1 - \lambda A} b_0 = \frac{S_{m-1}(\lambda) b_0 + S_{m-2}(\lambda)\lambda b_1 + \cdots + S_0 \lambda^{m-1} b_{m-1}}{S_m(\lambda)}. \tag{17}$$

[5] In order to prevent this paper from becoming too lengthy the analytical details of the present investigation are kept to a minimum, and in a few places the reader is requested to interpolate the missing steps.

If we compare this solution with the earlier solution (4), we notice that the expansion (17) may be conceived as a modified form of the Liouville-Neumann series, because it is composed of the same kind of terms, the difference being only that we *weight* the terms $\lambda^k b_k$ by the weight factors

$$w_k = \frac{S_{m-k-1}(\lambda)}{S_m(\lambda)}, \tag{18}$$

instead of taking them all with the uniform weight factor 1. This weighting has the beneficial effect that the series *terminates* after m terms, instead of going on endlessly. The weight factors w_i are very near to 1 for small λ but become more and more important as λ increases. The weighting makes the series convergent for *all* values of λ.

The remarkable feature of the expansion (17) is its *complete generality*. No matter how defective the matrix A may be, and no matter how the vector b_0 was chosen, the expansion (17) is *always* valid, provided only that we interpret it properly. In particular we have to bear in mind that there will always be m polynomials $S_k(\lambda)$, even though every $S_k(\lambda)$ may not be of degree k, due to the vanishing or the higher coefficients. For example, it could happen that

$$G(x) = x^m, \tag{19}$$

so that

$$S_m(\lambda) = 1 + 0\lambda + 0\lambda^2 + \ . \ . \ . \ + 0\lambda^m, \tag{20}$$

$$S_k(\lambda) = 1 + 0\lambda + 0\lambda^2 + \ . \ . \ . \ + 0\lambda^k, \tag{21}$$

and the formula (17) gives:

$$y = b_0 + \lambda b_1 + \ . \ . \ . \ + \lambda^{m-1} b_{m-1}. \tag{22}$$

IV. Solution of the Eigenvalue Problem

The Liouville-Neumann series cannot give the solution of the eigenvalue problem, since the expansion becomes divergent as soon as the parameter λ reaches the lowest characteristic number λ_1. The Schmidt series cannot give the solution of the eigenvalue problem since it presupposes the knowledge of all the eigenvalues and eigenvectors of the matrix A. On the other hand, the expansion (17), which is based purely on iterations and yet remains valid for all λ, must contain implicitly the solution of the principal axis problem. Indeed, let us write the right side of (1) in the form

$$b = S_m(\lambda)\overline{b}. \tag{23}$$

UNIVERSITY OF MICHIGAN LIBRARIES

Then the expansion (17) loses its denominator and becomes:

$$y=S_{m-1}(\lambda)\overline{b}_0+S_{m-2}(\lambda)\lambda\overline{b}_1+\ \ldots\ +S_0\lambda^{m-1}\overline{b}_{m-1}. \quad (24)$$

We can now answer the question whether a solution of the homogeneous equation

$$y-\lambda Ay=0, \quad (25)$$

is possible without the identical vanishing of y. The expression (23) shows that b can vanish only under two circumstances; either the vector $\overline{b}$, or the scalar $S_m(\lambda)$ must vanish. Since the former possibility leads to an identically vanishing y, only the latter possibility is of interest. This gives the following condition for the parameter λ.

$$S_m(\lambda)=0. \quad (26)$$

The roots of this equation give us the characteristic values $\lambda=\lambda_i$, whereas the solution (24) yields the characteristic solutions, or eigenvalues, or principal axes of the matrix A:

$$u_i=S_{m-1}(\lambda_i)b_0+S_{m-2}(\lambda_i)\lambda_i b_1+\ \ldots\ +S_0\lambda_i^{m-1}b_{m-1}. \quad (27)$$

It is a remarkable fact that although the vector b_0 was chosen entirely freely, the particular linear combination (27) of the iterated vectors has *invariant significance*, except for an undetermined factor of proportionality that remains free, in view of the linearity of the defining equation (25). That undetermined factor may come out even as zero, i. e., a certain axis may not be represented in the trial vector b_0 at all. This explains why the order of the polynomial $S_m(\lambda)$ need not be necessarily equal to n. The trial vector b_0 may not give us *all* the principal axes of A. What we can say with assurance, however, is that all the roots of $S_m(\lambda)$ are true characteristic values of A, and all the u_i obtained by the formula (24) are true characteristic vectors, even if we did not obtain the *complete* solution of the eigenvalue problem. The discrepancy between the order m of the polynomial $G(\mu)$ and the order n of the characteristic equation

$$F(\mu)=\begin{vmatrix} a_{11}-\mu & \ldots & a_{1n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{n1} & & a_{nn}-\mu \end{vmatrix}=0 \quad (28)$$

will be the subject of the discussion in the next section.

Instead of substituting into the formula (27), we can also obtain the principal axes u_i by a numerically simpler process, applying synthetic division. By synthetic division we generate the polynomials:

$$\frac{G(x)}{x-\mu_i}=x^{m-1}+g_1^i x^{m-2}+\ \ldots\ +g_{m-1}^i. \quad (29)$$

We then replace x^j by b_j and obtain:

$$u_i=b_{m-1}+g_1^i b_{m-2}+\ \ldots\ +g_{m-1}^i b_0. \quad (30)$$

The proof follows immediately from the equation

$$(A-\mu_i)u_i-G(A)\cdot b_0=0. \quad (31)$$

V. The Problem of Missing Axes

Let us assume that we start with an arbitrary "trial vector" b_0 and obtain by successive iterations the sequence:

$$b_0, b_1, b_2, \ldots, b_n. \quad (32)$$

Similarly we start with the trial vector b_0^* and obtain by iterating with the transposed matrix A^* the adjoint sequence:

$$b_0^*, b_1^*, b_2^*, \ldots, b_n^*. \quad (33)$$

Let us now form the following set of "basic scalars":

$$c_{i+k}=b_i\cdot b_k^*=b_k\cdot b_i^*. \quad (34)$$

The remarkable fact holds that these scalars depend only on the *sum* of the two subscripts i and k; e. g.

$$b_{k-1}b_{k+1}^*=b_{k-1}^*b_{k+1}=b_k b_k^*. \quad (35)$$

This gives a powerful numerical check of the iteration scheme since a discrepancy between the two sides of (35) (beyond the limits of the rounding errors) would indicate an error in the calculation of b_{k+1} or b_{k+1}^* if the sequence up to b_k and b_k^* has been checked before.

Let us assume for the sake of the present argument that A is a nondefective matrix, and let us analyze the vector b_0 in terms of the eigenvectors u_i, while b_0^* will be analyzed in terms of the adjoint vectors u_i^*:

$$b_0=\tau_1 u_1+\tau_2 u_2+\ \ldots\ +\tau_n u_n, \quad (36)$$

$$b_0^*=\tau_1^* u_1^*+\tau_2^* u_2^*+\ \ldots\ +\tau_n^* u_n^*. \quad (37)$$

Then the scalars c_i become:

$$c_i=\rho_1\mu_1^i+\rho_2\mu_2^i+\ \dots\ +\rho_n\mu_n^i, \tag{38}$$

with

$$\rho_k=\tau_k\tau_k^*. \tag{39}$$

The problem of obtaining the μ_i from the c_i is the problem of "weighted moments", which can be solved as follows: Assuming that none of the ρ_i vanish and that all the λ_i are distinct, we establish a linear relation between $n+1$ consecutive c_i, of the following form:

$$\left.\begin{aligned} &c_0\eta_0+c_1\eta_1+\ \dots\ +c_{n-1}\eta_{n-1}+c_n=0\\ &c_1\eta_0+c_2\eta_1+\ \dots\ +c_n\eta_{n-1}+c_{n+1}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_n\eta_0+c_{n+1}\eta_1+\ \dots\ +c_{2n-1}\eta_{n-1}+c_{2n}=0 \end{aligned}\right\} \tag{40}$$

Then the definition of the c_i shows directly that the set (40) demands

$$F(\mu_i)=0, \tag{41}$$

where

$$F(x)=\eta_0+\eta_1x+\ \dots\ +\eta_{n-1}x^{n-1}+x^n. \tag{42}$$

Hence, by solving the recurrent set (40) with the help of a "progressive algorithm", displayed in the next section, we can obtain the coefficients of the characteristic polynomial (42), whose roots give the eigenvalues μ_i.

Under the given restricting conditions none of the μ_i roots has been lost, and we could actually establish the full characteristic equation (41). It can happen, however, that b_0 is orthogonal to some axis u_i^*, and it is equally possible that b_0^* is orthogonal to some axis u_k. In that case τ_i and τ_k^* drop out of the expansions (36) and (37) and consequently the expansion (38) lacks both ρ_i and ρ_k. This means that the scalars c_j are unable to provide all the μ_j, since μ_i and μ_k are missing. The characteristic equation (41) cannot be fully established under these circumstances.

The deficiency was here caused by an unsuitable choice of the vectors b_0 and b_0^*; it is removable by a better choice of the trial vectors. However, we can have another situation where the deficiency goes deeper and is not removable by *any* choice of the trial vectors. This happens if the μ_i roots of the characteristic equation are not all distinct. The expansion (38) shows that two equal roots λ_i and λ_k cannot be separated since they behave exactly as one single root with a double amplitude. Generally, the weighted moments c_i can never show whether or not there are multiple roots, because the multiple roots behave like single roots. Consequently, in the case of multiple eigenvalues the linear relation between the c_i will not be of the n^{th} but of a lower order. If the number of distinct roots is m, then the relation (40) will appear in the following form:

$$\left.\begin{aligned} &c_0\eta_0+c_1\eta_1+\ \dots\ +c_{m-1}\eta_{m-1}+c_m=0\\ &c_1\eta_0+c_2\eta_1+\ \dots\ +c_m\eta_{m-1}+c_{m+1}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_m\eta_0+c_{m+1}\eta_1+\ \dots\ +c_{2m-1}\eta_{m-1}+c_{2m}=0 \end{aligned}\right\}. \tag{43}$$

Once more we can establish the polynomial

$$G(x)=\eta_0+\eta_1x+\ \dots\ +\eta_{m-1}x^{m-1}+x^m, \tag{44}$$

but this polynomial is now of only m^{th} order and factors into the m root factors

$$(x-\mu_1)(x-\mu_2)\ \dots\ (x-\mu_m), \tag{45}$$

where all the μ_i are distinct. After obtaining all the roots of the polynomial (44) we can now construct by synthetic division the polynomials:

$$\frac{G(x)}{x-\mu_k}=x^{m-1}+g_1^kx^{m-2}+\ \dots\ +g_{m-1}^k, \tag{46}$$

and replacing x^j by b_j we obtain the principal axes of both A and A^*:

$$\left.\begin{aligned} u_k&=b_{m-1}+g_1^kb_{m-2}+\ \dots\ +g_{m-1}^kb_0,\\ u_k^*&=b_{m-1}^*+g_1^kb_{m-2}^*+\ \dots\ +g_{m-1}^kb_0^*. \end{aligned}\right\} \tag{47}$$

This gives a partial solution of the principal axis problem, inasmuch as each multiple root contributed only one axis. Moreover, we cannot tell from our solution which one of the roots is single and which one multiple, nor can the degree of multiplicity be established. In order to get further information, we have to change our trial vectors and go through the iteration scheme once more. We now substitute into the formulae (47) again and can immediately localize all the single roots by the fact that the vectors u_j associated with these roots do not change (apart from a proportionality factor), whereas the u_k belonging to double roots will generally change their direction.

UNIVERSITY OF MICHIGAN LIBRARIES

A proper linear combination of the new u_k^*, and the previous u_k establishes the second axis associated with the double eigenvalue μ_k; we put

$$u_k^1 = u_k \qquad u_k^2 = u_k + \gamma u_k$$

$$u_k^{1*} = u_k^* \qquad u_k^{2*} = u_k^* + \gamma^* u_k^*$$

The factors γ and γ^* are determined by the conditions that the vectors u_k^1 and u_k^2 have to be biorthogonal to the vectors u_k^{1*} and u_k^{2*}. In the case of triple roots a third trial is demanded, and so on.

An interesting contrast to this behavior of multiple roots associated with nondefective matrices is provided by the behavior of multiple roots associated with defective matrices. A defective eigenvalue is always a multiple eigenvalue, but here the multiplicity is not caused by the collapse of two very near eigenvalues, but by the multiplicity of the elementary divisor. This comes into evidence in the polynomial $G(x)$ by giving a root factor of higher than first order. Whenever the polynomial $G(x)$ reveals a multiple root, we can tell in advance that the matrix A is defective in these roots, and the multiplicity of the root establishes the degree of deficiency.

It will be revealing to demonstrate these conditions with the help of a matrix that combines all the different types of irregularities that may be encountered in working with arbitrary matrices. Let us analyze the following matrix of sixth order:

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvalue 2 is the only regular eigenvalue of this matrix. The matrix is "singular", because the determinant of the coefficients is zero. This, however, is irrelevant from the viewpoint of the eigenvalue problem, since the eigenvalue "zero" is just as good as any other eigenvalue. More important is the fact that the eigenvalue zero is a *double root* of the characteristic equation. The remaining three roots of the characteristic equation are all 1. This 1 is thus a triple root of the characteristic equation; at the same time the matrix has a double deficiency in this root, because the elementary divisor associated with this root is cubic. The matrix possesses only four independent principal axes.

What will the polynomial $G(x)$ become in the case of this matrix? The regular eigenvalue 2 must give the root factor $x-2$. The regular eigenvalue 0 has the multiplicity 2 but is reduced to the single eigenvalue 0 and thus contributes the factor x. The deficient eigenvalue 1 has the multiplicity 3 but also double defectiveness. Hence, it must contribute the root factor $(x-1)^3$. We can thus predict that the polynomial $G(x)$ will come out as follows.

$$G(x) = x(x-2)(x-1)^3 = x^5 - 5x^4 + 9x^3 - 7x^2 + 2x.$$

Let us verify this numerically. As a trial vector we choose

$$b_0 = b_0^* = 1, 1, 1, 1, 1, 1.$$

The successive iterations yield the following:

$b_0 =$	1	1	1	1	1	1
$b_1 =$	6	5	1	2	0	0
$b_2 =$	19	9	1	4	0	0
$b_3 =$	40	13	1	8	0	0
$b_4 =$	69	17	1	16	0	0
$b_5 =$	106	21	1	32	0	0
$b_6 =$	151	25	1	64	0	0
$b_0^* =$	1	1	1	1	1	1
$b_1^* =$	1	3	8	2	0	0
$b_2^* =$	1	5	23	4	0	0
$b_3^* =$	1	7	46	8	0	0
$b_4^* =$	1	9	77	16	0	0
$b_5^* =$	1	11	116	32	0	0
$b_6^* =$	1	13	163	64	0	0

We now construct the c_i by dotting b_0 with the b_i^* (or b_0^* with the b_i); we continue by dotting b_1 with $b_1^*, \ldots b_6^*$ (or b_6^* with $b_1, \ldots b_6$). This gives the following set of $2n+1=13$ basic scalars.

$c_j = 6, 14, 33, 62, 103, 160, 241, 362, 555, 884, 1477, 2590, 4735.$

The application of the progressive algorithm of section VI to these b_i yields $G(x)$ in the predicted form. We now obtain by synthetic divisions:

$$\frac{G(x)}{x-1}=x^4-4x^3+5x^2-2x,$$

$$\frac{G(x)}{x-2}=x^4-3x^3+3x^2-x,$$

$$\frac{G(x)}{x}=x^4-5x^4+9x^2-7x+2.$$

Inverting these polynomials we obtain the matrix

$$\begin{pmatrix} 0 & -2 & 5 & -4 & 1 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 \\ 2 & -7 & 9 & -5 & 1 & 0 \end{pmatrix}$$

The product of this matrix with the iteration matrix B (omitting the last row b_6) yields three principal axes u_i; similarly the product of the same matrix with the iteration matrix B^* yields the three adjoint axes u_i^*:

$$\begin{aligned} u(1) &= -8 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \\ u(2) &= \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad 0 \\ u(0) &= \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 \quad 2 \end{aligned}$$

$$\begin{aligned} u^*(1) &= \quad 0 \quad 0 -8 \quad 0 \quad 0 \quad 0 \\ u^*(2) &= \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad 0 \\ u^*(0) &= \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 \quad 2 \end{aligned}$$

Since $G(x)$ is of only fifth order, while the order of the characteristic equation is 6, we know that one of the axes is still missing. We cannot decide a priori whether the missing axis is caused by the duplicity of the eigenvalue 0, 1, or 2.[9] However, a repetition of the iteration with the trial vectors

$$b_0=b_0^*=1,\ 1,\ 1,\ 1,\ 1,\ 0,$$

causes a change in the row $u(0)$ and $u^*(0)$ only. This designates the eigenvalue $u=0$ as the double root. The process of biorthogonalization finally yields: [10]

$$\begin{aligned} u_1(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \\ u_2(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1 \end{aligned}$$

$$\begin{aligned} u_1^*(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \\ u_2^*(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1 \end{aligned}$$

VI. The Progressive Algorithm for the Construction of the Characteristic Polynomial G(x)

The crucial point in our discussions was the establishment of a linear relation between a certain b_m and the previous iterated vectors. This relation leads to the characteristic polynominal $G(x)$, whose roots $G(\mu_i)=0$ yield the eigenvalues μ_i. Then by synthetic division we can immediately obtain those particular linear combinations of the iterated vectors b_i, which give us the eigenvectors (principal axes) of the matrix A.

We do not know in advance in what relation the order m of the polynomial $G(x)$ will be to the order n of the matrix A. Accidental deficiencies of the trial vectors b_0, b_0^*, and the presence of multiple eigenvalues in A can diminish m to any value between 1 and n. For this reason we will follow a systematic procedure that generates $G(x)$ gradually, going through all degrees from 1 to m. The procedure comes automatically to a halt when the proper m has been reached.

Our final goal is to solve the recurrent set of equations:

$$\left.\begin{aligned} &c_0\eta_0+c_1\eta_1+\ \ldots\ +c_m=0 \\ &\cdot \\ &\cdot \\ &\cdot \\ &c_m\eta_0+c_{m+1}\eta_1+\ \ldots\ +c_{2m}=0 \end{aligned}\right\}. \qquad (48)$$

This is only possible if the determinant of this homogeneous set vanishes:

$$\begin{vmatrix} c_0 & c_1 & \ldots & c_m \\ c_1 & c_2 & \ldots & c_{m+1} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ c_m & c_{m+1} & \ldots & c_{2m} \end{vmatrix} =0. \qquad (49)$$

[9] We have in mind the general case and overlook the fact that the oversimplified nature of the example makes the decision trivial.

[10] The reader is urged to carry through a similar analysis with the same matrix, but changing the 0, 0 diagonal elements of the rows 5 and 6 to 1, 0, and to 1, 1.

UNIVERSITY OF MICHIGAN LIBRARIES

Before reaching this goal, however, we can certainly solve for any $k<m$ the following inhomogeneous set (the upper k is meant as a superscript):

$$\left.\begin{aligned} &c_0\eta_0^k+c_1\eta_1^k+\ \ldots\ +c_k=0\\ &c_1\eta_0^k+c_2\eta_1^k+\ \ldots\ +c_{k+1}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_k\eta_0^k+c_{k+1}\eta_1^k+\ \ldots\ +c_{2k}=h_k \end{aligned}\right\}. \tag{50}$$

The freedom of h_k in the last equation removes the overdetermination of the set (48). The proper m will be reached as soon as h_m turns out to be zero.

Now a recurrent set of equations has certain algebraic properties that are not shared by other linear systems. In particular, there exists a *recurrence relation* between the solutions of three consecutive sets of the type (50). This greatly facilitates the method of solution, through the application of a systematic recurrence scheme that will now be developed.

We consider the system (50) and assume that we possess the solution up to a definite k. Then we will show how this solution may be utilized for the construction of the next solution, which belongs to the order $k+1$.

Our scheme becomes greatly simplified if we consider an additional set of equations that omits the first equation of (50) but adds one more equation at the end:

$$\left.\begin{aligned} &c_1\overline{\eta}_1^k+c_2\overline{\eta}_2^k+\cdots+c_{k+1}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_{k+1}\overline{\eta}_1^k+c_{k+2}\overline{\eta}_2^k+\cdots+c_{2k+1}=\overline{h}_{k+1} \end{aligned}\right\} \tag{51}$$

Let us now multiply the set (50) by the factor

$$q_k=-\frac{\overline{h}_{k+1}}{h_k}, \tag{52}$$

and add the set (51). We get a new set of equations that can be written down in the form:

$$\left.\begin{aligned} &c_0\eta_0^{k+1}+c_1\eta_1^{k+1}+\cdots+c_{k+1}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_k\eta_0^{k+1}+c_{k+1}\eta_1^{k+1}+\cdots+c_{2k+1}=0 \end{aligned}\right\} \tag{53}$$

provided that we put

$$\left.\begin{aligned} &\eta_0^{k+1}=q_k\eta_0^k\\ &\eta_1^{k+1}=q_k\eta_1^k+\overline{\eta}_1^k\\ &\cdot\\ &\cdot\\ &\cdot\\ &\eta_k^{k+1}=q_k\cdot 1+\overline{\eta}_k^k \end{aligned}\right\}. \tag{54}$$

We now evaluate the scalar

$$c_{k+1}\eta_0^{k+1}+\cdots+c_{2k+1}\eta_k^{k+1}+c_{2k+2}=h_{k+1} \tag{55}$$

which is added to (53) as the last equation.

What we have accomplished is that a proper linear combination of the solutions η_i^k and $\overline{\eta}_i^k$ provided us with the next solution η_i^{k+1}. But now exactly the same procedure can be utilized to obtain $\overline{\eta}_i^{k+1}$ on the basis of $\overline{\eta}_i^k$ and η_i^{k+1}.

For this purpose we multiply the set (51) by

$$\overline{q}_k=-\frac{h_{k+1}}{\overline{h}_{k+1}}, \tag{56}$$

and add the set (53), completed by (55) but omitting the first equation.
This gives:

$$\left.\begin{aligned} &c_1\overline{\eta}_1^{k+1}+c_2\overline{\eta}_2^{k+1}+\cdots+c_{k+2}=0\\ &\cdot\\ &\cdot\\ &\cdot\\ &c_{k+1}\overline{\eta}_1^{k+1}+c_{k+2}\overline{\eta}_2^{k+1}+\cdots+c_{2k+2}=0 \end{aligned}\right\} \tag{57}$$

provided that we put:

$$\left.\begin{aligned} &\overline{\eta}_1^{k+1}=\overline{q}_k\,\overline{\eta}_1^k+\eta_0^{k+1}\\ &\overline{\eta}_2^{k+1}=\overline{q}_k\,\overline{\eta}_2^k+\eta_1^{k+1}\\ &\cdot\\ &\cdot\\ &\cdot\\ &\overline{\eta}_{k+1}^{k+1}=\overline{q}_k\cdot 1+\eta_k^{k+1} \end{aligned}\right\} \tag{58}$$

Once more we evaluate the scalar

$$c_{k+2}\overline{\eta}_1^{k+1}+\cdots+c_{2k+3}=\overline{h}_{k+2}, \tag{59}$$

which is added to (57) as the last equation.

This analytical procedure can be translated into an elegant geometrical arrangement that generates the successive solutions η_i^k and $\overline{\eta}_i^k$ in successive columns. The resulting algorithm is best explained with the help of a numerical example.

For this purpose we choose the eigenvalue problem of an intentionally over-simplified matrix, since our aim is not to show the power of the

method but the nature of the algorithm, which leads to the establishment of the characteristic equation. The limitation of the method due to the accumulation of rounding errors will be discussed in the next section.

Let the given matrix be:

$$\begin{pmatrix} 13 & 5 & -23 \\ 4 & 0 & -4 \\ 7 & 3 & -13 \end{pmatrix}$$

We iterate with the trial vector $b_0=1, 0, 0$, and obtain:

1	0	0
13	4	7
28	24	12
208	64	112

We transpose the matrix and iterate with the trial vector $b_0^*=1, 0, 0$, obtaining:

1	0	0
13	5	−23
28	−4	−20
208	80	−368

We dot the first row and the last row with the opposing matrix and obtain the basic scalars c_i as follows:[11]

1, 13, 28, 208, 448, 3328, 7168.

These numbers are written down in a column, and the scheme displayed below is obtained.

	0	0.5	1	1.5	2	2.5	3
$h_i=$	1	13	−141	147.6923075	−163.4042569	0	0
$q'=$	−13	10.84615384	1.047463173	1.106382990	0		
1	1						
13		1	−13				
28			1	−2.15384616	−13.617021249		
208				1	−1.106382987	−16.000000004	0
448					1	0.00000003	−16
3328						1	0
7168							1

(60)

Instead of distinguishing between the η_i and $\bar{\eta}_i$ solutions we use a uniform procedure but mark the successive columns alternately as "full" and "half columns"; thus we number the successive columns as zero, one-half, one, . . . The scheme has to end at a *full* column, and the end is marked by the vanishing of the corresponding "head-number" h_i. In our scheme the head-number is zero already at the half-column 2.5, but the scheme cannot end here, and thus we continue to the column 3, whose head-number becomes once more 0, and then the scheme is finished. The last column gives the polynomial $G(x)$, starting with the diagonal term and proceeding upward:

$$G(x)=1\cdot x^3+0\cdot x^2-16x+0,$$
$$=x^3-16x.$$

The head-numbers h_i are always obtained by dotting the column below with the basic column c_i; e. g., at the head of the column 2 we find the number −163.4042569. This number was obtained by the following cumulative multiplication:

$$448\cdot 1+208\cdot(-1.106382987)+28\cdot(-13.617021249).$$

The numbers q_i represent the negative ratio of two consecutive h_i numbers:

$$q_i=-\frac{h_{i+\frac{1}{2}}}{h_i},$$

[11] Instead of iterating with A and A^* n times, we can also iterate with A alone $2n$ times. Any of the columns of the iteration matrix can now be chosen as c_i numbers since these columns correspond to a dotting of the iteration matrix with $b_0^*=1, 0, 0, \ldots$, respectively $0, 1, 0, 0, \ldots$; $0, 0, 1, 0, 0, \ldots$; and so on. The transposed matrix is not used here at all. E. C. Bower of the Douglas Aircraft Co. points out to the author that from the machine viewpoint a uniform iteration scheme of $2n$ iterations is preferable to a divided scheme of $n+n$ iterations. The divided scheme has the advantage of less accumulation of rounding errors and more powerful checks on the successive iterations. The uniform scheme has the advantage that more than one column is at our disposal. Accidental deficiencies of the b_0^* vector can thus be eliminated, by repeating the algorithm with a different column. (For this purpose it is of advantage to start with the trial vector $b_0=1, 1, 1, \ldots 1$.) In the case of a *symmetric* matrix it is evident that after n iterations the basic scalars should be formed, instead of continuing with n more iterations.

UNIVERSITY OF MICHIGAN LIBRARIES

e. g., $q_{1.5}=1.106382990$ was obtained by the following division:

$$\frac{-(-163.4042569)}{147.6923075}.$$

The scheme grows as follows. As soon as a certain column C_i is completed, we evaluate the associated head-number h_i; this provides us with the previous q-number $q_{i-\frac{1}{2}}$. We now construct the next column $C_{i+\frac{1}{2}}$ by the following operation. We multiply the column $C_{i-\frac{1}{2}}$ by the constant $q_{i-\frac{1}{2}}$ and add the column C_i:

$$C_{i+\frac{1}{2}}=q_{i-\frac{1}{2}}\cdot C_{i-\frac{1}{2}}+C_i.$$

However, the result of this operation is shifted down by one element; e. g. in constructing the column 2.5 the result of the operation

$$1.10638299\cdot(-2.15384616)+(-13.617021249)=-16,$$

is not put in the row where the operation occurred, but shifted down to the next row below.

The unfilled spaces of the scheme are all "zero".

The outstanding feature of this algorithm is that *it can never come to premature grief*, provided only that the first two c-numbers, c_0 and c_1, are different from zero. Division by zero cannot occur since the scheme comes to an end anyway as soon as the head-number zero appears in one of the full columns.

Also of interest is the fact that the products of the head-numbers associated with the full columns give us the successive recurrent determinants of the c_i; e.g., the determinants

$$1 \quad \begin{vmatrix} 1 & 13 \\ 13 & 28 \end{vmatrix} \quad \begin{vmatrix} 1 & 13 & 28 \\ 13 & 28 & 208 \\ 28 & 208 & 448 \end{vmatrix}$$

and

$$\begin{vmatrix} 1 & 13 & 28 & 208 \\ 13 & 28 & 208 & 448 \\ 28 & 208 & 448 & 3328 \\ 208 & 448 & 3328 & 7168 \end{vmatrix}$$

are given by the successive products

$$1,\ 1\cdot(-141)=-141,\ (-141)\cdot(-163.4042569)=23040,\text{ and }23040\cdot 0=0,$$

Similarly the products of the head-numbers of the half-columns give us similar determinants, but omitting c_0 from the sequence of c-numbers. In the example above the determinants

$$13, \quad \begin{vmatrix} 13 & 28 \\ 28 & 208 \end{vmatrix} \quad \begin{vmatrix} 13 & 28 & 208 \\ 28 & 208 & 448 \\ 208 & 448 & 3328 \end{vmatrix}$$

are given by the products:

$$13,\ 13\cdot 147.6923075=1920,\ 1920\cdot 0=0$$

The purpose of the algorithm (60) was to generate the coefficients of the basic identity that exists between the iterated vectors b_i. This identity finds expression in the vanishing of the polynomial $G(x)$:

$$G(x)=0 \tag{61}$$

The roots of this algebraic equation give us the eigenvalues of the matrix A. In our example we get the cubic equation

$$x^3-16x=0,$$

which has the three roots

$$\mu_1=0,\ \mu_2=4,\ \mu_3=-4. \tag{62}$$

These are the eigenvalues of our matrix. In order to obtain the associated eigenvectors, we divide $G(x)$ by the root factors:

$$\frac{G(x)}{x}=x^2-16 \quad \frac{G(x)}{x-4}=x^2+4x \quad \frac{G(x)}{x+4}=x^2-4x.$$

This gives, replacing x^k by b_k:

$$u(0)=-16b_0+b_2,$$

$$u(4)=4b_1+b_2,$$

$$u(-4)=-4b_1+b_2.$$

Consequently, if the matrix

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix}$$

is multiplied by the matrix of the b_i (omitting b_3), we obtain the three eigenvectors u_i:

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 13 & 4 & 7 \\ 28 & 24 & 12 \end{pmatrix} = \begin{pmatrix} 12 & 24 & 12 \\ 80 & 40 & 40 \\ -24 & 8 & -16 \end{pmatrix} = \begin{pmatrix} u(0) \\ u(4) \\ u(-4) \end{pmatrix}.$$

If the same matrix is multiplied by the matrix of the transposed iterations b_i^* (omitting b_3^*), we obtain the three adjoint eigenvectors u_i^*:

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 13 & 5 & -23 \\ 28 & -4 & -20 \end{pmatrix} = \begin{pmatrix} 12 & -4 & -20 \\ 80 & 16 & -112 \\ -24 & -24 & 72 \end{pmatrix} = \begin{pmatrix} u^*(0) \\ u^*(4) \\ u^*(-4) \end{pmatrix}.$$

The solution of the entire eigenvalue problem is thus accomplished.

VII. The Method of Minimized Iterations

In principle, the previous discussions give a complete solution of the eigenvalue problem. We have found a systematic algorithm for the generation of the characteristic polynomial $G(\mu)$. The roots of this polynomial gave the eigenvalues of the matrix A. Then the process of synthetic division established the associated eigenvectors. Accidental deficiencies were possible but could be eliminated by additional trials.

As a matter of fact, however, the "progressive algorithm" of the last section has its serious limitations if large matrices are involved. Let us assume that there is considerable "dispersion" among the eigenvalues, which means that the ratio of the largest to the smallest eigenvalue is fairly large. Then the successive iterations will grossly increase the gap, and after a few iterations the small eigenvalues will be practically drowned out. Let us assume, e. g., that we have a 12 by 12 matrix, which requires 12 iterations for the generation of the characteristic equation. The relatively mild ratio of 10:1 as the "spread" of the eigenvalues is after 12 iterations increased to the ratio 10^{12}:1, which means that we can never get through with the iteration scheme because the rounding errors make all iterations beyond the eighth entirely valueless.

As an actual example, taken from a physical situation, let us consider four eigenvalues, which are distributed as follows:

$$1, \quad 5, \quad 50, \quad 2000.$$

Let us assume, furthermore, that we start with a trial vector that contains the four eigenvectors in the ratio of the eigenvalues, i. e., the eigenvalue 2000 dominates with the amplitude 2000, compared with the amplitude of the eigenvalue 1. After one iteration the amplitude ratio is increased to $4 \cdot 10^6$, after two iterations to $8 \cdot 10^9$. The later iterations can give us no new information, since they practically repeat the second iteration, multiplied every time by the factor 2000. The small eigenvalues 1 and 5 are practically obliterated and cannot be rescued, except by an excessive accuracy that is far beyond the limitations of the customary digital machines.

We will now develop a modification of the customary iteration technique that dispenses with this difficulty. The modified scheme eliminates the rapid accumulation of rounding errors, which under ordinary circumstances destroys the value of high order iterations. The new technique prevents the large eigenvalues from monopolizing the scene. It protects the small eigenvalues by constantly balancing the distribution of amplitudes in the most equitable fashion.

As an illustrative example, let us apply this method of "minimized iterations" to the above-mentioned dispersion problem. If the largest amplitude is normalized to 1, then the initial distribution of amplitudes is characterized as follows:

$$0.0005, \quad 0.0025, \quad 0.025, \quad 1.$$

Now, while an ordinary iteration would make this distribution still more extreme, the method of minimized iterations changes the distribution of amplitudes as follows:

$$0.0205 \quad 0.1023 \quad 1 \quad -0.0253.$$

We see that it is now the *third* eigenvector that gets a large weight factor, whereas the fourth eigenvector is almost completely in the background.

A repetition of the scheme brings about the following new distribution:

$$0.2184 \quad 1 \quad -0.1068 \quad 0.0000.$$

UNIVERSITY OF MICHIGAN LIBRARIES

It is now the *second* eigenvector that gets the strongest emphasis.

The next repetition yields:

$$1 \quad -0.2181 \quad 0.0018 \quad 0.0000,$$

and we see that the weight shifted over to the *smallest* eigenvalue.

After giving a chance to each eigenvalue, the scheme is exhausted, since we have all the information we need. Consequently the next minimized iteration yields an identical *vanishing* of the next vector, thus bringing the scheme to its natural conclusion.

In order to expose the principle of minimized iterations, let us first consider the case of *symmetric* matrices:

$$A^* = A. \tag{63}$$

Moreover, let us agree that the multiplication of a vector b by the matrix A shall be denoted by a prime:

$$Ab = b'. \tag{64}$$

Now our aim is to establish a linear identity between the iterated vectors. We cannot expect that this identity will come into being right from the beginning. Yet we can *approach* this identity right from the beginning by choosing that linear combination of the iterated vectors b_0' and b_0 which makes the amplitude of the new vector as small as possible. Hence, we want to choose as our new vector b_1 the following combination:

$$b_1 = b_0' - \alpha_0 b_0, \tag{65}$$

where α_0 is determined by the condition that

$$(b_0' - \alpha_0 b_0)^2 = \text{minimum}. \tag{66}$$

This gives

$$\alpha_0 = \frac{b_0' b_0}{b_0^2}. \tag{67}$$

Notice that

$$b_1 \cdot b_0 = 0, \tag{68}$$

i.e. the new vector b_1 is orthogonal to the original vector b_0.

We now continue our process. From b_1 we proceed to b_2 by choosing the linear combination

$$b_2 = b_1' - \alpha_1 b_1 - \beta_0 b_0, \tag{69}$$

and once more α_1 and β_0 are determined by the condition that b_2^2 shall become as small as possible. This gives

$$\alpha_1 = \frac{b_1' b_1}{b_1^2} \qquad \beta_0 = \frac{b_1' b_0}{b_0^2}. \tag{70}$$

A good check of the iteration b_1' is provided by the condition

$$b_1' b_0 = b_1 b_0' = b_1^2. \tag{71}$$

Hence, the numerator of β_0 has to agree with the denominator of α_1.

The new vector b_2 is orthogonal to both b_0 and b_1.

This scheme can obviously be continued. The most remarkable feature of this successive minimization process is, however, that the best linear combination *never includes more than three terms*. If we form b_3, we would think that we should put

$$b_3 = b_2' - \alpha_2 b_2 - \beta_1 b_1 - \gamma_0 b_0. \tag{72}$$

But actually, in view of the orthogonality of b_2 to the previous vectors, we get

$$\gamma_0 = \frac{b_2' b_0}{b_0^2} = \frac{b_2 b_0'}{b_0^2} = 0. \tag{73}$$

Hence, every new step of the minimization process requires only two correction terms.

By this process a succession of orthogonal vectors is generated:[12]

$$b_0, \; b_1, \; b_2, \; \ldots, \; b_{m-1}, \tag{74}$$

until the identity relation becomes exact, which means that

$$b_m = 0. \tag{75}$$

If the matrix A is not symmetric, then we modify our procedure as follows: We operate simultaneously with A and A^*. The operations are the same as before, with the only difference that the dot products are always formed between two *opposing* vectors. The scheme is indicated as follows:

$$\begin{array}{cc} b_0 & b_0^* \\ b_1 = b_0' - \alpha_0 b_0 & b_1^* = b_0^{*\prime} - \alpha_0 b_0^* \end{array}$$

$$\alpha_0 = \frac{b_0' b_0^*}{b_0 b_0^*} = \frac{b_0^{*\prime} b_0}{b_0^* b_0}$$

[12] The idea of the successive orthogonalization of a set of vectors was probably first employed by O. Szász, in connection with a determinant theorem of Hadamard; cf. Math. és phys. lapok (in Hungarian) **19**, 221 to 227, (1910). The method found later numerous important applications.

$$b_2=b_1'-\alpha_1 b_1-\beta_0 b_0 \qquad b_2^*=b_1^{*\prime}-\alpha_1 b_1^*-\beta_0 b_0^*$$

$$\alpha_1=\frac{b_1' b_1^*}{b_1 b_1^*}=\frac{b_1^{*\prime} b_1}{b_1^* b_1} \tag{76}$$

$$\beta_0=\frac{b_1' b_0^*}{b_0 b_0^*}=\frac{b_1^{*\prime} b_0}{b_0^* b_0}=\frac{b_1^* b_1}{b_0^* b_0}$$

$$b_2=b_2'-\alpha_2 b_2-\beta_1 b_1 \qquad b_3^*=b_2^*-\alpha_2 b_2^*-\beta_1 b_1^*, \text{ etc.}$$

Operationally the prime indicates multiplication by the matrix A. Hence, the succession of b_i vectors represents in fact a successive *set of polynomials*. Replacing A by the more familiar letter x, we have:

$$\left.\begin{aligned} b_0&=1\cdot b_0\\ b_1&=(x-\alpha_0)b_0\\ b_2&=(x-\alpha_1)b_1-\beta_0 b_0\\ b_3&=(x-\alpha_2)b_2-\beta_1 b_1\\ &\cdots\cdots\\ b_m&=(x-\alpha_{m-1})b_{m-1}-\beta_{m-2}b_{m-2}=0 \end{aligned}\right\} \tag{77}$$

This gradual generation of the characteristic polynomial $G(x)$ is in complete harmony with the procedure of the "progressive algorithm", discussed in section VI. In fact, *the successive polynomials of the set* (77) *are identical with the polynomials found in the full columns of the progressive algorithm* (60). This explains the existence of the recurrence relation

$$p_{n+1}(x)=(x-\alpha_n)p_n(x)-\beta_{n-1}p_{n-1}(x), \tag{78}$$

without additional γ, δ, . . . terms. The existence of such a relation is a characteristic feature of the *recurrent* set of equations that are at the basis of the entire development.

Although the new scheme goes basically through the same steps as the previously discussed "progressive algorithm", it is in an incomparably stronger position concerning rounding errors. Apart from the fact that the rounding errors do not accumulate, we can effectively counteract their influence by constantly checking the mutual orthogonality of the gradually evolving vectors b_i and b_i^*. Any lack of orthogonality, caused by rounding errors, can immediately be corrected by the addition of a small correction term.[13] By this procedure the orthogonality of the generated vector system does not come gradually out of gear.

However, quite apart from the numerical advantages, the biorthogonality of the vectors b_i and b_i^* has further appeal because *it imitates the behavior of the principal axes*. This is analytically an eminently valuable fact that makes the transition from the iterated vectors to the principal axes a simple and strongly convergent process.

In order to see the method in actual operation, let us apply it to the simple example of section VI. Here the matrix A is of third order, and thus we have to construct the vectors b_0, b_1, b_2, b_3, and the corresponding adjoint vectors. We obtain the following results:

$$b_0=\ 1\quad 0\quad 0 \qquad b_0^*=\ 1\quad 0\quad 0$$

$$b_0'=13\quad 4\quad 7 \qquad b_0^{*\prime}=13\quad 5\quad -23$$

$$\alpha_0=\frac{1\cdot 13}{1}=13$$

$$b_1=\ 0\quad 4\quad 7 \qquad b_1^*=\ 0\quad 5\quad -23$$

$$b_1'=-141\quad -28\quad -79 \qquad b_1^{*\prime}=-141\quad -69\quad 279$$

$$\alpha_1=\frac{1677}{-141} \qquad \beta_0=\frac{-141}{1}=-141$$

$$=-11.89361702$$

$$b_2=0,\ 19.57446808,\ 4.25531914$$

$$b_2^*=0,\ -9.53191490,\ 5.44680854$$

$$b_2'=0,\ -17.02127656,\ 3.40425542$$

$$b_2^{*\prime}=0,\ 16.34042562,\ -32.68085142$$

$$\alpha_2=\frac{180.78768715}{-163.4042553} \qquad \beta_1=\frac{-163.40425746}{-141}$$

$$=-1.106382981 \qquad =1.158895443$$

$$b_3=0,\quad 0,\quad 0 \qquad b_3^*=0,\quad 0,\quad 0$$

The associated polynomials become:

$$p_0=1$$

$$p_1(x)=x-13$$

$$p_2(x)=(x+11.89361702)(x-13)+141$$

$$=x^2-1.10638298x-13.61702126$$

$$p_3(x)=(x+1.106382981)(x^2+10638298x-13.61702126)-1.158895443(x-13)$$

$$=x^3-16x$$

[13] See section IX, eq 90.

UNIVERSITY OF MICHIGAN LIBRARIES

Comparison with the scheme (60) shows that the coefficients of these very same polynominals appear in the full columns 0, 1, 2, 3, of the progressive algorithm of section VI.

VIII. Solution of the Eigenvalue Problem by the Method of Minimized Iterations

The biorthogonal property of the vector system b_i, b_i^* leads to an explicit solution of the eigenvalue problem, in terms of the vectors b_i. Let us first assume that the matrix A is of the nondefective type, and let us analyze the vectors b_i in terms of the eigenvectors u_i. The method by which the vectors b_i were generated, yields directly the relation

$$b_i=p_i(\mu_1)u_1+p_i(\mu_2)u_2+\ \ldots\ +p_i(\mu_m)u_m. \tag{79}$$

If this relation is dotted with u_k^*, we obtain, in view of the mutual orthogonality of the two sets of axes:

$$b_i\cdot u_k^*=p_i(\mu_k)u_k\cdot u_k^*. \tag{80}$$

Let us now reverse the process and expand the u_i in terms of the b_i:

$$u_1=\alpha_{10}b_0+\alpha_{11}b_1+\ \ldots\ +\alpha_{1m-1}b_{m-1}. \tag{81}$$

The dotting by b_k^* yields:

$$\alpha_{1k}=\frac{u_1\cdot b_k^*}{b_k\cdot b_k^*}. \tag{82}$$

Let us denote the "norm" of b_k by σ_k:

$$\sigma_k=b_k\cdot b_k^*, \tag{83}$$

while the norm of u_k will be left arbitrary. Then the expansion (81) becomes:

$$u_1=\frac{b_0}{\sigma_0}+p_1(\mu_1)\frac{b_1}{\sigma_1}+p_2(\mu_1)\frac{b_2}{\sigma_2}+\ \ldots\ +p_{m-1}(\mu_1)\frac{b_{m-1}}{\sigma_{m-1}}. \tag{84}$$

This expansion contains the solution of the principal axis problem. The eigenvectors u_i are generated in terms of the vectors b_i, which are the successive vectors of the process of minimized iterations. The expansion (84) takes the place of the previous "S-expansion" (27), which solved the eigenvector problem in terms of the customary process of iteration.

The adjoint axes are obtained in identical fashion:

$$u_1^*=\frac{b_0^*}{\sigma_0}+p_1(\mu_1)\frac{b_1^*}{\sigma_1}+p_2(\mu_1)\frac{b_2^*}{\sigma_2}+\ \ldots\ +p_{m-1}(\mu_1)\frac{b_{m-1}^*}{\sigma_{m-1}}. \tag{85}$$

The expansion (84) remains valid even in the case of defective matrices. The only difference is that the number of principal axes becomes less than n since a multiple root μ_j, if substituted into (84) and (85) cannot contribute more than one principal axis.[14] However, a defective matrix actually possesses less than n pairs of principal axes, and the above expansions give the general solution of the problem.

An interesting alternative of the expansion (84) arises if we go back to the original Fredholm problem and request a solution in terms of the minimized vectors b_i, rather than the simply iterated vectors of the expansion (17). One method would be to make use of the Schmidt series (8), expressing the u_i of that series in terms of the b_i, according to the expansion (84). However, the Schmidt series holds for nondefective matrices only, while we know that a solution must exist for any kind of matrices.

Hence, we prefer to proceed in a somewhat different fashion. We expand y directly in terms of the vectors b_i:

$$y=y_0b_0+y_1b_1+\ .\ .\ .+y_{m-1}b_{m-1}. \tag{86}$$

We substitute this expansion into the equation (1), replacing b_k' by

$$b_k'=b_{k+1}+\alpha_kb_k+\beta_{k-1}b_{k-1}. \tag{87}$$

Then we compare coefficients on both sides of the equation. The result can be described as follows:

Let us reverse the sequence of the α_i-coefficients and let us do the same with the β_i-coefficients. Hence, we define

$$\begin{aligned}
&\bar\alpha_0=\alpha_{m-1}\\
&\bar\alpha_1=\alpha_{m-2}\qquad \bar\beta_0=\beta_{m-2}\\
&\ .\qquad\qquad\ .\\
&\ .\qquad\qquad\ .\\
&\ .\qquad\qquad\ .\\
&\bar\alpha_{m-1}=\alpha_0\qquad \bar\beta_{m-2}=\beta_0
\end{aligned} \tag{88}$$

---

[14] The reader is urged to carry through the process of minimized iterations and evaluation of the principal axes for the defective matrix

$$\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$$

which has only one pair of principal axes; (choose the trial vector in the form $b_0=b_0^*=1, 1, 1$).

We now construct the following "reversed" set of polynomials:

$$\begin{aligned}
\overline{p}_0 &= 1 \\
\overline{p}_1(x) &= x-\overline{\alpha}_0 \\
\overline{p}_2(x) &= (x-\overline{\alpha}_1)\overline{p}_1(x)-\overline{\beta}_0 \qquad (89)\\
&\vdots \\
\overline{p}_m(x) &= (x-\overline{\alpha}_{m-1})\overline{p}_{m-1}(x)-\overline{\beta}_{m-2}\overline{p}_{m-2}(x) \\
&= G(x)
\end{aligned}$$

Then the solution of the Fredholm problem (1) is given by the following expansion:

$$y=\frac{\mu}{G(\mu)}\,[b_{m-1}+\overline{p}_1(\mu)b_{m-2}(u)b_{m-2}+\ \ldots\ +\overline{p}_{m-1}(u)b_0], \qquad (90)$$

where we have put

$$\mu=\frac{1}{\lambda}. \qquad (91)$$

The expansion (90) is *completely general* and remains valid, no matter how the vector b_0 of the right side was given, and how regular or irregular the matrix A may be. The only condition to be satisfied is that the vector b_0^*—while otherwise chosen arbitrarily—shall be free of accidental deficiencies, i. e., b_0^* shall not be orthogonal to some u_k if b_0 is not simultaneously orthogonal to u_k^*.

The expansion (90) leads once more to a solution of the eigenvector problem, this time obtained with the help of the "reversed" polynomials $\overline{p}_i(x)$:

$$u_i=b_{m-1}+\overline{p}_1(\mu_i)b_{m-2}+\ \ldots\ +\overline{p}_{m-1}(\mu_i)b_0. \qquad (92)$$

The expansions (92) and (84) actually coincide—except for a factor of proportionality—due to algebraic reasons.

In order to see a numerical example for this solution of the eigenvalue problem let us return once more to the simple problem of section VI. The minimized b_i and b_i^* vectors associated with this matrix were given at the end of section VII. together with the associated polynomials $p_i(x)$. We now construct the *reversed* polynomials $\overline{p}_i(x)$. For this purpose we tabulate the α_i and β_i:

13	−141
−11.89361702	1.158895443
−1.106382981	

We reverse the sequence of this tabulation:

−1.106382981	
−11.89361702	1.58895443
13	−141

and construct in succession:

$$\begin{aligned}
\overline{p}_0 &= 1 \\
\overline{p}_1(x) &= x+1.106382981 \\
\overline{p}_2(x) &= (x+11.8361702)\overline{p}_1(x)-1.58895443 \\
&= x^2+13x+12 \\
\overline{p}_3(x) &= (x-13)\overline{p}_2(x)+141 \\
&= x^3-16x
\end{aligned}$$

The last polynomial is identical with $p_3(x)=G(x)$. The zeros of this polynomial are:

$$\mu_1=0, \qquad \mu_2=4, \qquad \mu_3=-4;$$

substituting these values into $\overline{p}_2(\mu)$, $\overline{p}_1(\mu)$, $\overline{p}_0$ we obtain the matrix:

$$\begin{pmatrix} 12 & 1.10638298 & 1 \\ 80 & 5.10638298 & 1 \\ -24 & -2.89361702 & 1 \end{pmatrix}$$

The product of this matrix with the matrix of the b_i vectors gives the three principal axes u_i:

$$\begin{pmatrix} 12 & 1.10638298 & 1 \\ 80 & 5.10638298 & 1 \\ -24 & -2.89361702 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 7 \\ 0 & 19.57446808 & 4.25531914 \end{pmatrix}=\begin{pmatrix} 12 & 24 & 12 \\ 80 & 40 & 40 \\ -24 & 8 & -16 \end{pmatrix}=\begin{pmatrix} u(0) \\ u(4) \\ u(-4) \end{pmatrix}$$

UNIVERSITY OF MICHIGAN LIBRARIES

in complete agreement with the previous result of section VI, but now obtained by an entirely different method. If the b-matrix is replaced by the b^*-matrix, the adjoint axes $u^*(0)$, $u^*(4)$. $u^*(-4)$ are obtained.

IX. The Lateral Vibrations of a Bar

In order to study the power of the method in connection with a vibration problem of large dispersion, the elastic vibrations of a bar were investigated. The bar was clamped at one end and free at the other. Moreover, the cross section of the bar changes in the middle, as shown in figure 1.

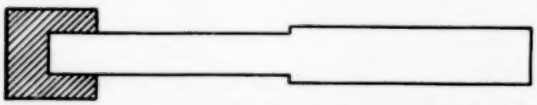

FIGURE 1. *Elastic bar, clamped at one end, free at the other; the moment of inertia makes a jump in the middle in the ratio 1:2.*

The change of the cross section was such that the moment of inertia jumped from the value 1 to 2. The differential equation that describes the vibrations of such a bar is the following fourth-order equation:

$$\frac{d^2}{dx^2}\left[k(x)\frac{d^2y}{dx^2}\right]=\lambda y, \tag{93}$$

with the boundary conditions:

$$\left.\begin{aligned} y(0)&=0 \qquad y''(l)=0 \\ y'(0)&=0 \qquad y'''(l)=0 \end{aligned}\right\} \tag{94}$$

and

$$k(x)=1\left(0\leqq x<\frac{l}{2}\right),\ k(x)=2\left(\frac{l}{2}<x\leqq l\right) \tag{95}$$

The differential operator d/dx was replaced by the difference operator $\Delta/\Delta x$, with $\Delta x=1$. The length of the bar was chosen as $l=13$, thus leading to a 12 by 12 matrix; (since $y(0)=y(1)=0$).

The first step was the inversion of the matrix. This was easily accomplished, since a matrix that is composed of a narrow band around the diagonal can be inverted with little labor. The eigenvalues μ_i of the inverted matrix are the *reciprocals* of the original λ_i:

$$\mu_k=\frac{1}{\lambda_i}. \tag{96}$$

The general theory has shown that the iteration scheme applied to an arbitrary matrix automatically yields a biorthogonal set of vectors b_i and b_i^*; they can be conceived as the building blocks from which the entire set of principal axes may be generated. In the present problem dissipative forces are absent, which makes the matrix A symmetric and the problem self-adjoint. Hence,

$$b_i=b_i^*, \tag{97}$$

and we get through with a single set of iterations.

Now the general procedure would demand that we go through 12 minimized iterations before the stage $b_{12}=0$ is attained. However, the study of a system with high dispersion has shown that in such a system the method of minimized iterations practically separates the various vibrational modes, starting with the highest eigenvalue and descending systematically to the lower eigenvalues, provided that we employ a trial vector b_0 which weights the eigenvectors approximately according to the associated eigenvalues. In the present problem the trial vector 1, 0, 0 . . ., was not used directly but iterated with the matrix A, and then iterated again. The thus obtained vector b_0^* was employed as the b_0 of the minimized iteration scheme.

The strong grading of the successive eigenvectors has the consequence that in k minimized iterations essentially only the highest k vibrational modes will come into evidence. This is of eminent practical value, since it allows us to dispense with the calculation of the very low eigenvalues (i. e., very high frequencies since we speak of the eigenvalues of the *inverted* matrix), which are often of little physical interest, and also of little mathematical interest, in view of the fact that the replacing of the d operator by the Δ operator becomes in the realm of high frequencies more and more damaging.

Whether the isolation actually takes place or not can be tested with the help of the $p_i(x)$ polynomials, which accompany the iteration scheme. The order of these polynomials constantly increases by 1. The correct eigenvalues of the matrix A are obtained by evaluating the zeros of the *last* polynomial $p_m(x)=0$. What actually happens, however, is that the zeros of the polynomials $p_i(x)$ do not change much from the beginning. If the dispersion is strong, then each new polynomial basically adds one more root but corrects the higher roots by only small amounts. It is

thus well possible that the series of largest roots in which we are primarily interested is practically established with sufficient accuracy after a few iterations. Then we can stop, since the later iterations will change the obtained values by negligible amounts. The same can be said about the vibrational modes associated with these roots.

This consideration suggests the following successive procedure for the approximate determination of the eigenvalues and eigenvectors (vibrational modes) of a matrix. As the minimization scheme proceeds and we constantly obtain newer and newer polynomials $p_i(x)$, we handle the last polynominal obtained as if it were the final polynomial $p_m(x)$. We evaluate the roots of this polynomial and compare them with the previous roots. Those roots that change by negligible amounts are already in their final form.

A similar procedure holds for the evaluation of the eigenvectors u_i. Here the biorthogonality of the vectors b_i and b_i^* —which is reduced to simple orthogonality in the case of a symmetric matrix— is of very great help. Let us assume that the length of the vectors b_i is normalized to 1, by replacing b_i by $b_i/\sqrt{\sigma_i}$. Then the expansions (84) and (85) show that the following matrix must be an orthogonal—although in the diagonal terms not normalized—matrix:

$$\begin{pmatrix} \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_1)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_1)}{\sqrt{\sigma_2}} & \cdots & \frac{p_{m-1}(\mu_1)}{\sqrt{\sigma_{m-1}}} \\ \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_2)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_2)}{\sqrt{\sigma_2}} & \cdots & \frac{p_{m-1}(\mu_2)}{\sqrt{\sigma_{m-1}}} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_m)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_m)}{\sqrt{\sigma_2}} & \cdots & \frac{p_{m-1}(\mu_m)}{\sqrt{\sigma_{m-1}}} \end{pmatrix} \tag{98}$$

The dot-product of any two rows of this matrix must come out as zero—thus providing us with a powerful check on the construction of the p_i polynomials and the correctness of the roots μ_i, which are the roots of the equation $p_m(\mu)=0$.[15] In the case of strong dispersion, the transformation matrix (98) is essentially reduced to the diagonal terms and one term to the right and to the left of the diagonal; i. e., the eigenvector u_k is essentially a linear combination of three b-vectors only, viz., b_{k-2}, b_{k-1}, and b_k.

These general conditions are well demonstrated by the tabulation of the final results of the abovementioned bar problem. The minimized iterations were carried through up to $m=6$. On the basis of these iterations the first six eigenvalues and the first five vibrational modes of the clamped-free bar were evaluated. The iterations were constantly watched for orthogonality. After obtaining a certain b_i, this b_i was immediately dotted with all the previous b_j. If a certain dot-product $b_i \cdot b_j$ came out as noticeably different from zero, the correction term

$$\epsilon_{ij} = -\frac{b_i \cdot b_j}{b_j^2} b_j, \tag{99}$$

was added to b_i, thus compensating for the influence of rounding errors. By this procedure the 10 significant figure accuracy of the calculations was constantly maintained.[16]

The roots of the successive polynomials $p_i(x)$ are tabulated as follows:

μ_1	μ_2	μ_3
2256.926071		
.943939	48.1610705	
.943939	.2037755	5.272311428
.943939	.2037825	.355958260
.943939	.2037825	.356269794
.943939	.2037825	.356269980

μ_4	μ_5	μ_6
1.513923859		
.582259337	0.546327303	
.5829955952	.891117817	0.2498132719

[15] Due to reasons of algebra, the orthogonality of the matrix (98) holds not only for the final m but for *any value of* m.

[16] In a control experiment that imitated the conditions of the vibrating bar but with a more regular matrix, the results were analytically predictable and the computational results open to an exact check. This example vividly demonstrated the astonishing degree of noninterference of orthogonal vectors. The spread of the eigenvalues was 1 : 3200. The trial vector b_0 strongly over-emphasized the largest eigenvalue, containing the lowest and the highest eigenvectors with an amplitude ratio of $1 : 10^8$; (this means that if the vector of the smallest eigenvalue were drawn with the length of 1 in., the vector of the largest eigenvalue, perpendicular to the previous one, would span the distance from Los Angeles to Chicago). The slightest inclination between the two vectors would fatally injure the chances of the smallest eigenvalue. When the entire computation scheme was finished, the analytically required eigenvalues were computed and the comparison made. ***The entire set of eigenvalues, including the last, agreed with a maximum error of 2 units in the ninth significant figure,*** thus demonstrating that the method is completely free of the cumulation of rounding errors. The author is indebted to Miss Fannie M. Gordon, of the National Bureau of Standards, for the eminently careful and skilled performance of the computing operations.

UNIVERSITY OF MICHIGAN LIBRARIES

The successive orthogonal transformation matrices (98) show likewise strong convergence. We tabulate here only the last computed transformation matrix (rounded off to four decimal places), which expresses the first six eigenvectors $u_1, \ldots, u_6$ in terms of the first six normalized $b_i/\sqrt{\sigma_i}$ vectors, making use of the roots of $p_6(u)=0$. The diagonal elements are normalized to 1:

$$\begin{bmatrix} 1 & .0028 & 0 & 0 & 0 & 0 \\ -.0028 & 1 & .0316 & .0004 & 0 & 0 \\ 0 & -.0316 & 1 & .1497 & .0081 & 0 \\ 0 & .0044 & -.1520 & 1 & .2693 & .0249 \\ 0 & -.0010 & .0335 & -.2793 & 1 & .4033 \\ 0 & .0002 & -.0087 & .0779 & -.3816 & 1 \end{bmatrix}$$

We notice how quickly the elements fall off to zero as soon as we are beyond one element to the right and one to the left of the main diagonal. The orthogonal reference system of the b_i and the orthogonal reference system of the u_i are thus in close vicinity of each other.

The obtained five vibrational modes $u_1 \ldots, u_5$ (u_6 being omitted since the lack of the neighbour on the right side of the diagonal makes the approximation unreliable) are plotted graphically in figure 2.

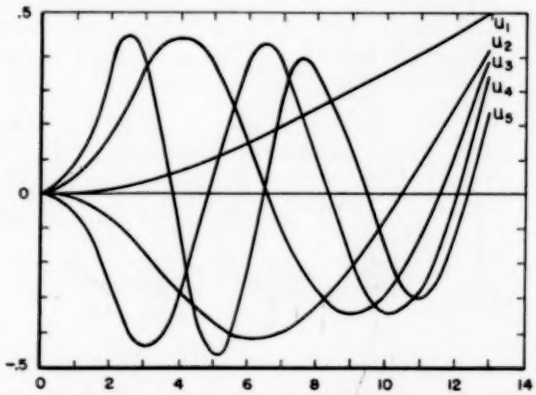

FIGURE 2. *The first five lateral vibrational modes of the elastic bar, shown in figure 1.*

$\mu_1=2256.94$; $\mu_2=48.20$; $\mu_3=5.36$; $\mu_4=1.58$; $\mu_5=0.59$

X. The Eigenvalue Problem of Linear Integral Operators

The methods and results of the past sections can now be applied to the realm of continuous operators. The kernel of an integral equation can be conceived as a matrix of infinite order, which may be approximated to any degree of accuracy by a matrix of high but finite order. One method of treating an integral equation is to replace it by an ordinary matrix equation of sufficiently high order. This procedure is, from the numerical standpoint, frequently the most satisfactory one. However, we can design methods for the solution of integral equations that obtain the solution by purely analytical tools, on the basis of an infinite convergent expansion, such as the Schmidt series, for example. The method we are going to discuss belongs to the latter type. We will find an expansion that is based on the same kind of iterative integrations as the Liouville-Neumann series, but avoiding the convergence difficulties of that expansion. The expansion we are going to develop *converges under all circumstances* and gives the solution of any Fredholm type of integral equation, no matter how defective the kernel of that integral equation may be.[17]

Let us first go back to our earlier method of solving the Fredholm problem. The solution was obtained in form of the S-expansion (17). The difficulty with this solution is that it is based on the linear identity that can be established between the iterated vectors b_i. That identity is generally of the order n; if n grows to infinity, we have to obtain an identity of infinite order before our solution can be constructed. That, however, cannot be done without the proper adjustments.

The later attempt, based on the method of minimized iterations, employs more adequate principles. We have seen that to any matrix, A, a biorthogonal set of vectors b_i and b_i^* can be constructed by successive minimizations. The set is uniquely determined as soon as the first trial vectors b_0 and b_0^* are given. In the case of the inhomogeneous equation (1) the right side b may be chosen as the trial vector b_0, while b_0^* is still arbitrary.

[17] The Volterra type of integral equations, which have no eigenvalues and eigensolutions, are thus included in our general considerations.

The construction of these two sets of vectors is quite independent of the order n of the matrix. If the matrix becomes an integral operator, the b_i and b_i^* vectors are transformed into a biorthogonal set of functions:

$$\begin{matrix} \phi_0(x), & \phi_1(x), & \phi_2(x), & \dots \\ \phi_0^*(x), & \phi_1^*(x), & \phi_2^*(x), & \dots \end{matrix} \tag{100}$$

which are generally present in infinite number. The process of minimized iterations assigns to *any* integral operator such a set, after $\phi_0(x)$ and $\phi_0^*(x)$ have been chosen.

Another important feature of the process of minimized iterations was the appearance of a successive set of polynomials $p_i(\mu)$, tied together by the recurrence relation

$$p_{i+1}(\mu)=(\mu-\alpha_i)p_i(\mu)-\beta_{i-1}p_{i-1}(\mu). \tag{101}$$

This is again entirely independent of the order n of the matrix A and remains true even if the matrix A is replaced by a Fredholm kernel $K(x,\xi)$.

We can now proceed as follows. We stop at an arbitrary $p_m(x)$ and form the reversed set of polynomials $\overline{p}_i(\mu)$, defined by the process (88). Then we construct the expansion:

$$y_m(x)=\frac{\mu}{p_m(\mu)}\left[\phi_{m-1}(x)+\overline{p}_1(\mu)\phi_{m-2}(x)+\dots+\overline{p}_{m-1}(\mu)\phi_0(x)\right]. \tag{102}$$

This gives a successive approximation process that converges well to the solution of the Fredholm integral equation

$$y(x)-\lambda Ky(x)=\phi_0(x). \tag{103}$$

In other words:

$$y(x)=\lim_{m\to\infty} y_m(x). \tag{104}$$

By the same token we can obtain all the eigenvalues and eigensolutions of the kernel $K(x,\xi)$, if such solutions exist. For this purpose we obtain the roots μ_i of the polynomial $p_m(\mu)$ by solving the algebraic equation

$$p_m(\mu)=0. \tag{105}$$

The exact eigenvalues μ_i of the integral operator $K(x,\xi)$ are obtained by the limit process:

$$\lim_{m\to\infty} p_m(\mu_i)=0, \tag{106}$$

where the largest root is called μ_1, and the subsequent roots are arranged according to their absolute values. The corresponding eigenfunctions are given by the following infinite expansion:

$$\mu_i(x)=\lim_{m\to\infty}\left[\frac{\phi_0(x)}{\sigma_0}+p_1(\mu_i)\frac{\phi_1(x)}{\sigma_1}+\dots+p_{m-1}(\mu_i)\frac{\phi_{m-1}(x)}{\sigma_{m-1}}\right], \tag{107}$$

where μ_i is the i^{th} root of the polynomial $p_m(\mu)$.[18]

As a trial function $\phi_0(x)$ we may choose for example

$$\phi_0=\text{const.}=1. \tag{108}$$

However, the convergence is greatly speeded up if we first apply the operator K to this function, and possibly iterate even once more. In other words, we should choose $\phi_0=K\cdot 1$, or even $\phi_0=K^2\cdot 1$ as the basic trial function of the expansion (107).

We consider two particularly interesting examples that are well able to illustrate the nature of the successive approximation process here discussed.

(A) *The vibrating membrane.* In the problem of the vibrating membrane we encounter the following self-adjoint differential operator:

$$-\frac{d}{dx}(xy'). \tag{109}$$

This leads to the eigenvalue problem

$$-\frac{d}{dx}(xy')=\lambda y\,(0\leqq x\leqq 1), \tag{110}$$

with the boundary condition

$$y(1)=0. \tag{111}$$

The solution of the differential equation (110) is

$$y=J_0(2\sqrt{\lambda x}), \tag{112}$$

where $J_0(x)$ is the Bessel-function of the order zero. The boundary condition (111) requires that λ shall be chosen as follows:

$$\lambda_i=\frac{\xi_i^2}{4}, \tag{113}$$

where ξ_i are the zeros of $J_0(x)$.

[18] This expansion is not of the nature of the Neumann series, because the coefficients of the expansion are not rigid but constantly changing with the number m of approximating terms.

UNIVERSITY OF MICHIGAN LIBRARIES

Now the Green's function of the differential equation (110) changes the differential operator (109) to the inverse operator, which is an integral operator of the nature of a symmetric Fredholm kernel function $K(x, \xi)$. Our problem is to obtain the eigenvalues and eigenfunctions of this kernel.

If we start out with the function $\phi_0(x)=1$, the operation $K\phi_0$ gives

$$1-x,$$

and repeating the operation we obtain

$$\frac{x^2}{4}-x+\frac{3}{4},$$

and so on. The successive iterations will be *polynomials* in x. Now the minimized iterations are merely some linear combinations of the ordinary iterations. Hence the orthogonal sequence $\phi_i(x)$ will become a sequence of polynomials of constantly increasing order, starting with the constant $\phi_0=1$. This singles out the $\phi_k(x)$ as the *Legendre polynomials* $P_k(x)$, but normalized to the range 0 to 1, instead of the customary range -1 to $+1$. The renormalization of the range transforms the polynomials $P_k(x)$ into Jacobi polynomials $G_k(p, q; x)$, with $p=q=1$ [8], which again are special cases of the Gaussian hypergeometric series $F(\alpha, \beta, \gamma; x)$, in the sense of $F(k+1, -k, 1; x)$; hence, we get:

$$\left.\begin{aligned}
\phi_0 &= 1\\
\phi_1(x) &= 1-2x\\
\phi_2(x) &= 1-6x+6x^2\\
\phi_3(x) &= 1-12x+30x^2-20x^3\\
&\cdots\cdots\cdots\cdots\cdots
\end{aligned}\right\} \tag{114}$$

The associated polynomials $p_i(x)$ can be obtained on the basis of the relation:

$$K\phi_m=\phi_{m+1}+\alpha_m\phi_m+\beta_{m-1}\phi_{m-1}. \tag{115}$$

This gives:

$$\left.\begin{aligned}
p_0 &= 1\\
p_1(x) &= 2x-1\\
p_2(x) &= 24x^2-18x+1\\
p_3(x) &= 720x^3-600x^2+72x-1
\end{aligned}\right\} \tag{116}$$

In order to obtain the general recurrence relation for these polynomials it is preferable to follow the example of the algorithm (60) and introduce the "half-columns" in addition to the full columns. Hence, we define a second set of polynomials $q_k(x)$ and set up the following recurrence relations:

$$\left.\begin{aligned}
p_n(x) &= nxq_{n-1}(x)-p_{n-1}(x)\\
q_n(x) &= 2(2n+1)p_n(x)-q_{n-1}(x)
\end{aligned}\right\} \tag{117}$$

We thus obtain, starting with $p_0=1$ and $q_0=2$, and using successive recursions:

$$\begin{aligned}
p_0(x) &= 1 & q_0(x) &= 2\\
p_1(x) &= 2x-1 & q_1(x) &= 12x-8\\
p_2(x) &= 24x^2-18x+1 & q_2(x) &= 240x^2-192x+18\\
p_3(x) &= 720x^3-600x^2+72x-1 & q_3(x) &= 10080x^3-8640x^2+1200x-32\\
&\cdots & &\cdots
\end{aligned} \tag{118}$$

The zeros of the $p_m(x)$ polynomials converge to the eigenvalues of our problem, but the convergence is somewhat slow since the original function $\phi_0=1$ does not satisfy the boundary conditions and thus does not suppress sufficiently the eigenfunctions of high order. The zeros of the q_i polynomials give quicker convergence. They are tabulated as follows, going up to $q_5(x)$:

μ_1	μ_2	μ_3	μ_4	μ_5
0.6667				
.69155	0.1084			
.69166016	.130242	.035241		
.6916602716	.1312564	.051130	.014842	
.6916602760	.13127115	.0532914	.025582	.00729
.6916602761	.13127123	.0534138	.028769	.01794

(119)

The last row contains the correct values of the eigenvalues, computed on the basis of (113):

$$\mu_i=\frac{4}{\xi_i^2}. \tag{120}$$

We notice the excellent convergence of the scheme.

The question of the eigenfunctions of our problem will not be discussed here.

(B) *The vibrating string. Even modes.* Another interesting example is provided by the vibrating string. The differential operator here is

$$-\frac{d^2}{dx^2}, \tag{121}$$

with the boundary conditions

$$y(\pm 1)=0. \tag{122}$$

The solution of the differential equation

$$-\frac{d^2}{dx^2}y=\lambda y, \tag{123}$$

under the given boundary conditions is:

$$y_i=\cos\ (2i+1)\ \frac{\pi}{2}\ x \quad \text{(even modes)}. \tag{124}$$

$$y_j=\sin\ j\pi x \qquad \text{(odd modes)}. \tag{125}$$

This gives the eigenvalues:

$$\lambda_i=\left(\frac{2i+1}{2}\ \pi\right)^2 \quad \text{(even modes)}, \tag{126}$$

and

$$\lambda_j=(j\pi)^2 \qquad \text{(odd modes)}. \tag{127}$$

If we start with the trial function $\phi_0=1$, we will get all the even vibrational modes of the string, while $\phi_0=x$ will give all the odd vibrational modes. We first choose the previous alternative.

Successive iterations give

$$\left.\begin{aligned} K\cdot 1&=\frac{x^2}{2}-\frac{1}{2}\\ K^2\cdot 1&=\frac{x^4}{24}-\frac{x^2}{4}+\frac{5}{24}\end{aligned}\right\} \tag{128}$$

and we notice that the minimized iterations will now become a sequence of *even* polynomials. The transformation $x^2=\xi$ shows that these polynomials are again Jacobi polynomials $G_k(p, q; x^2)$, but now $p=q=\frac{1}{2}$, and we obtain the hypergeometric functions $F(k+\frac{1}{2}, -k, \frac{1}{2}; x^2)$:

$$\left.\begin{aligned} \phi_0&=1\\ \phi_1(x)&=1-3x^2\\ \phi_2(x)&=3-30x^2+35x^4\\ \phi_3(x)&=5-105x^2+315x^4-231x^6\\ &\cdots\cdots\cdots\cdots\end{aligned}\right\} \tag{129}$$

Once more we can establish the associated polynomials $p_i(x)$, and the recurrence relation by which they can be generated. In the present case the recurrence relations come out as follows:

$$\left.\begin{aligned} p_n(x)&=(4n-1)xq_{n-1}(x)-p_{n-1}(x)\\ q_n(x)&=(4n+1)p_n(x)-q_{n-1}(x)\end{aligned}\right\} \tag{130}$$

starting with $p_0=1$, $q_0=1$. This yields

$$\left.\begin{aligned} p_0&=1\\ p_1(x)&=3x-1\\ p_2(x)&=105x^2-45x+1\\ p_3(x)&=10395x^3-4725x^2+210x-1\\ &\cdots\cdots\\ q_0&=1\\ q_1(x)&=15x-6\\ q_2(x)&=945x^2-420x+15\\ q_3(x)&=135135x^3-62370x^2+3150x-28\\ &\cdots\cdots\end{aligned}\right\} \tag{131}$$

The successive zeros of the $q_i(x)$ polynomials, up to $q_5(x)$, are tabulated as follows:

μ_1	μ_2	μ_3	μ_4	μ_5
0.40000				
.405059	0.02351			
.4052847733	.044856	0.011397		
.4052847346	.04503010	.015722	0.00455	
.4052847346	.0450316322	.016192	.00752	0.00216
.4052847346	.0450316371	.016211	.00827	.00500

The last row contains the correct eigenvalues, calculated from the formula:

$$\mu_i=\left(\frac{2}{2i+1}\ \frac{1}{\pi}\right)^2. \tag{132}$$

The convergence is again very conspicuous.

(C) *The vibrating string. Odd modes.* In the case of the odd modes of the vibrating string, the orthogonal functions of the minimized iterations are again related to the Jacobi polynomials $G_k(p, q; x)$, but now with $p=q=3/2$. Expressed in terms of the hypergeometric series we now get the polynomials of odd orders

$$xF\left(k+\frac{3}{2}, -k, \frac{3}{2}; x\right):$$

UNIVERSITY OF MICHIGAN LIBRARIES

$$\left.\begin{aligned}
\phi_0 &= x \\
\phi_1(x) &= 3x - 5x^3 \\
\phi_2(x) &= 15x - 70x^3 + 63x^5 \\
\phi_3(x) &= 35x - 315x^3 + 693x^5 - 429x^7 \\
&\cdots
\end{aligned}\right\} \quad (133)$$

The associated $p_n(x)$ polynomials are generated by the following recurrence relations:

$$\left.\begin{aligned}
p_n(x) &= (4n+1)xq_{n-1}(x) - p_{n-1}(x) \\
q_n(x) &= (4n+3)p_n(x) - q_{n-1}(x)
\end{aligned}\right\} \quad (134)$$

starting with $p_0 = 1$, $q_0 = 3$. We thus get

$$\left.\begin{aligned}
p_0 &= 1 \\
p_1(x) &= 15x - 1 \\
p_2(x) &= 945x^2 - 105x + 1 \\
p_3(x) &= 135135x^3 - 17325x^2 + 378x - 1 \\
&\cdots \\
q_0 &= 3 \\
q_1(x) &= 105x - 10 \\
q_2(x) &= 10395x^2 - 1260x + 21 \\
q_3(x) &= 2027025x^3 - 270270x^2 + 6930x - 36 \\
&\cdots
\end{aligned}\right\} \quad (135)$$

The table of the zeros of $q_i(x)$, up to $q_5(x)$, is given as follows:

μ_1	μ_2	μ_3	μ_4	μ_5
0.0952				
.10126	0.01995			
.10132106	.02500	0.00701		
.1013211836	.025323	.01098	0.00307	
.1013211836	.02533024	.011215	.00550	.00156
.1013211836	.025330296	.011258	.00633	.00405

The last row contains the correct eigenvalues calculated on the basis of the formula:

$$\mu_i = \frac{1}{i^2\pi^2}. \quad (136)$$

XI. The Eigenvalue Problem of Linear Differential Operators

Let $Dy(x)$ be a given linear differential operator, with given homogeneous boundary conditions of sufficient number to establish an eigenvalue problem. The problem of finding the eigenvalues and eigenfunctions of this operator is equivalent to the problem of the previous section in which the eigenvalue problem of linear integral operators was investigated. Let us assume that we know the Green's function $K(x,\xi)$ of the differential equation

$$Dy = \rho. \quad (137)$$

Then K is the reciprocal operator of D which possesses the same eigenfunctions (principal axes) as the operator D, whereas the eigenvalues of K are the reciprocals of the eigenvalues of D.

Hence, in principle the eigenvalue problem of differential operators needs no special investigation. Actually, however, the situation in most cases is far less simple. The assumption that we are in the possession of the Green's function associated with the differential equation (137) is often of purely theoretical significance. Even very simple differential operators have Green's functions, which are outside the limits of our analytical possibilities. Moreover, even if we do possess the integral operator K in closed form, it is still possible that the successive integrations needed for the construction of the successive orthogonal functions $\phi_1(x)$, $\phi_2(x)$, . . . go beyond our analytical facilities.

In view of this situation we ask the question whether we could not relax some of the practically too stringent demands of the general theory. We may lose somewhat in accuracy, but we may gain tremendously in analytical operations if we can replace some of the demands of the general theory by more simplified demands. The present section will show how that may actually be accomplished.

Leaving aside the method of minimized iterations, which was merely an additional tool in our general program, the basic principle of our entire investigation, if shaped to the realm of integral operators, may be formulated as follows:

We start out with a function $f_0(x)$, which may be chosen as $f_0(x) = 1$. We then form by iterated integrations a set of new functions

$$f_1(x) = Kf_0(x), f_2(x) = Kf_1(x), \ldots, f_m(x) = Kf_{m-1}(x). \quad (138)$$

Then we try to establish an approximate linear relation between these functions, as accurately as possible. For this purpose we made use of the method of least squares.

We notice that the general principle involves

two processes: (a) the construction of the iterated set (138); (b) the establishment of a close linear relation between them. It is the first process where the knowledge of the integral operator $K=D^{-1}$ is demanded. But let us observe that the relation between the successive f_i-functions can be stated in reverse order. We then get:

$$f_m(x), f_{m-1}(x)=Df_m(x), \ldots, f_0(x)=Df_1(x). \quad (139)$$

If we start with the function $f_m(x)$, then the successive functions of lower order can be formed with the help of the given D operator and we can completely dispense with the use of the Green's function.

Now the freedom of choosing $f_0(x)$ makes also $f_m(x)$ to some extent a free function. Yet, the successive functions $f_i(x)$ do not have the same degree of freedom. Although $f_0(x)$ need not satisfy the given boundary conditions, $f_1(x)$ of necessity satisfies these conditions, whereas $f_2(x)$ satisfies them more strongly, since not only $f_2(x)$ but even $Df_2(x)$ satisfies the given boundary conditions. Generally, we can say that an arbitrary $f_n(x)$ need not satisfy any definite differential or integral equation, but it is very restricted in the matter of boundary conditions: it has to satisfy the boundary conditions "to the n^{th} order." This means that not only $f_n(x)$ itself, but the whole series of functions:

$$f_n(x), Df_n(x), D^2f_n(x), \ldots D^{n-1}f_n(x), \quad (140)$$

have to satisfy the given boundary conditions.

To construct a function $f_n(x)$ with this property is not too difficult. We expand $f_n(x)$ into a linear set of powers, or periodic functions, or any other kind of functions we may find adequate to the given problem. The coefficients of this expansion will be determined by the boundary conditions that are satisfied by $f_n(x)$ and the iterated functions (140). This leads to the solution of linear equations. In fact, this process can be systematized to a regular recurrence scheme that avoids the accumulation of simultaneous linear equations, replacing them by a set of separated equations, each one involving but one unknown.

We have thus constructed our set (138), although in reverse order. We did not use any integrations, only the repeated application of the given differential operator D. The first phase of our problem is accomplished.

We now turn to the *second* phase of our program, viz., the establishment of an approximate linear relation between the iterated functions $f_i(x)$. The method of least squares is once more at our disposal. However, here again we might encounter the difficulty that the definite integrals demanded for the evaluation of the α_i and β_i are practically beyond our means. Once more we can simplify our task. The situation is similar to that of evaluating the coefficients of a Fourier series. The "best" coefficients, obtained by the method of least squares, demand the evaluation of a set of definite integrals. Yet we can get a practically equally close approximation of a function $f(x)$ by a finite trigonometric series $\overline{f(x)}$, if we use the method of "trigonometric interpolation." Instead of minimizing the mean square of the error $f(x)-\overline{f(x)}$, we make $\overline{f(x)}$ equal to $f(x)$ at a sufficient number of equidistant points. This leads not to integrations but to simple summations.

The present situation is quite analogous. To establish a linear relation between the $f_i(x)$ means that the last function $f_m(x)$ is to be approximated by a linear combination of the previous functions. Instead of using the method of least squares for this approximation we can use the much simpler method of *interpolation*, by establishing a linear relation between the successive $f_i(x)$ in as many equidistant points as we have coefficients at our disposal. For the sake of better convergence, it is preferable to omit $f_0(x)$—which does not satisfy the boundary conditions and thus contains the high vibrational modes too pronouncedly—and establish the linear relation only from $f_1(x)$ on. For example, if we constructed a trial function $f_3(x)$ which, together with the iterated $f_2(x)=Df_3(x)$ and doubly iterated $f_1(x)=D^2f(x)$ satisfies the given boundary conditions, then we can choose two points of the region, e.g., the two endpoints, where a linear relation of the form

$$f_3(x)+\alpha f_2(x)+\beta f_1(x)=0, \quad (141)$$

shall hold. This gives the characteristic polynomial $G(x)$ in the form

$$G(x)=x^2+\alpha x+\beta. \quad (142)$$

The two roots of this polynomial give us an approximate evaluation of the two highest μ_i, (or the two lowest λ_i), i.e., $\mu_1=1/\lambda_1$, and $\mu_2=1/\lambda_2$.

UNIVERSITY OF MICHIGAN LIBRARIES

whereas the corresponding eigensolutions are obtained by synthetic division:

$$\left.\begin{aligned}\frac{G(x)}{x-\mu_1}&=g_1'x+g_2'\\ \frac{G(x)}{x-\mu_2}&=g_1''x+g_2''\end{aligned}\right\} \tag{143}$$

which gives

$$\left.\begin{aligned}u_1(x)&=g_1'f_2(x)+g_1'f_1(x)\\ u_2(x)&=g_1''f_2(x)+g_2''f_1(x)\end{aligned}\right\} \tag{144}$$

(The last root and its eigenfunction are always considerably in error, and gives only rough indications.)

The remarkable feature of this method is that it *completely avoids any integrations*, requiring only the solution of a relatively small number of linear equations.

The following application of the method demonstrates its practical usefulness. The method was applied to obtain the first three eigenvalues of the lateral vibrations of a uniform bar, clamped at both ends. The given differential operator is here:

$$Dy=\frac{d^4y}{dx^4}$$

with the boundary conditions

$$y(\pm 1)=0 \qquad y'(\pm 1)=0.$$

Only the even modes were considered, expanding $y(x)$ into even powers of x. The approximations were carried out to first, second, and third order. The obtained eigenvalues are tabulated in the following table, the last row containing the correct eigenvalues given by Rayleigh [9]

μ_1	μ_2	μ_3
0.0323413		
.0319686	0.0007932	
.031963958	.0010875	0.0000788
0.031963936	0.0010946	0.0001795

We notice that the general convergence behavior of this method is exactly the same as that of the analytically more advanced, but practically much more cumbersome, method of minimized iterations.

XII. Differential Equations of Second Order; Milne's Method

If a linear differential equation of second order with two-end boundary conditions is changed into a difference equation and then handled as a matrix problem, singularly favorable conditions exist for the solution of the eigenvalue problem. The matrix of the corresponding difference equation contains only diagonal terms plus one term to the right and one to the left. If we now start to iterate with the trial vector

$$b_0=1,\ 0,\ 0,\ \ldots,\ 0, \tag{145}$$

we observe that the successive iterations grow by one element only, as indicated in the following scheme where the dots stand for the nonvanishing components:

$$\left.\begin{aligned}b_0 &= .\\ b_1 &= . \quad .\\ b_2 &= . \quad . \quad .\\ &\vdots\\ b_{n-1} &= . \quad . \quad . \quad . \quad . \quad . \quad .\\ b_n &= . \quad . \quad . \quad . \quad . \quad . \quad .\end{aligned}\right\} \tag{146}$$

Under these conditions the establishment of the linear identity between the iterated vectors is greatly simplified since it is available by a successive recurrence scheme. The coefficients of the equation

$$b_n+g_1b_{n-1}+g_2b_{n-2}+\ \ldots\ +g_nb_0=0, \tag{147}$$

are directly at our disposal, since the last column of the last two vectors gives g_1, then the previous column gives g_2, . . . , until finally the first column gives g_n. The construction of the basic polynomial $G(x)$ is thus accomplished and the eigenvalues λ_i[19] directly available by finding the roots of the equation $G(\lambda)=0$.

Professor W. E. Milne of the Oregon State College and the Institute for Numerical Analysis, applied the general theory to this problem, but with the following modification. Instead of iterat-

[19] We call the eigenvalues here λ_i, since the operator D is not inverted into an integral operator K.

ing with the given matrix A, Milne considers the regular vibration problem

$$\frac{\partial^2 u}{\partial t^2}+Du=0, \tag{148}$$

where the operator D has the following significance:[20]

$$Du=-\left[\frac{d^2}{dx^2}+p(x)\frac{d}{dx}+g(x)\right]u(x,t). \tag{149}$$

The differential equation (148) is now converted into a difference equation, with $\Delta x=\Delta t=h$. Then the value of $u(ih, jh)$ are determined by successive recursions, starting from the initial conditions

$$u(ih, 0)=1, 0, 0, 0, \ldots, 0, \tag{150}$$

and

$$u(ih, h)=u(ih, -h). \tag{151}$$

The linear identity between the $n+1$ vectors

$$u(ih, 0), u(ih, h), \ldots, u(ih, nh), \tag{152}$$

leads to a trigonometric equation for the characteristic frequencies ν_i, of the following form

$$\cos n\gamma_i h+A_{n-1}\cos(n-1)\nu_i h+\ldots+A_0=0. \tag{153}$$

We then put

$$\lambda_i=\nu_i^2. \tag{154}$$

On the other hand, the regular iteration method gives the eigenvalues $\bar{\lambda}_i$ of the operator Δu, defined in harmony with the operator Du but with the modification that the operation d/dx is replaced by the operation $\Delta/\Delta x$. The $\bar{\lambda}_i$ are in the following relation to the ν_i of eq 153 and 154:

$$\bar{\lambda}_i=\bar{\nu}_i^2=\left(\frac{\sin\frac{h}{2}\nu_i}{\frac{h}{2}}\right)^2. \tag{155}$$

It is of interest to see that the values (154) of Milne are much closer to the true eigenvalues than the values obtained by iterations. The values of Milne remain good even for high frequencies, whereas the iteration method gives gradually worse results; this is to be expected since the error committed by changing the differential equation to a difference equation must come into evidence with ever increasing force, as we proceed to the vibrational modes of higher order.

The following table illustrates the situation. It contains the results of one of Milne's examples; ("Example 1"). Here

$$D=-\left(\frac{d^2}{dx^2}+2\frac{d}{dx}\right), \tag{156}$$

with the boundary conditions:

$$u(0)=u(1)=0. \tag{157}$$

Moreover, h was chosen as $\frac{1}{8}$ and $n=7$. The column $\sqrt{\lambda_k}$ gives the correct frequencies, the column $\sqrt{\lambda_k^*}$ gives the frequencies obtained by Milne's method, while the column $\sqrt{\bar{\lambda}_k}$ gives the frequencies obtained by the iteration method.

k	$\sqrt{\lambda_k}$	$\sqrt{\lambda_k^*}$	$\sqrt{\bar{\lambda}_k}$
1	3.2969	3.2898	3.2667
2	6.3623	6.3457	6.1806
3	9.4777	9.4507	8.9107
4	12.6061	12.5664	11.3138
5	15.7398	15.6820	13.2891
6	18.8761	18.7870	14.7581
7	22.0139	21.8430	15.6629

Actually, it is purely a matter of computational preference whether we follow the one or the other scheme since there is a rigid relation between the two schemes. The frequencies ν_i obtained by Milne's method are in the following relation to the frequencies $\bar{\nu}_i$ obtained by the matrix iteration method:

$$\nu_i\frac{\sin\frac{h}{2}\nu_i}{\frac{h}{2}\nu_i}=\bar{\nu}_i \tag{158}$$

Hence, the results obtained by the one scheme can be translated into the results of the other scheme, and vice versa.

This raises the question, why it is so beneficial to transform the frequencies $\bar{\nu}_i$ of the Δu operator to the frequencies ν_i by the condition

$$\sin\frac{h}{2}\nu_i=\frac{h}{2}\bar{\nu}_i. \tag{159}$$

The answer is contained in the fact that the correction factor

$$\frac{\sin\frac{h}{2}\nu_i}{\frac{h}{2}\nu_i}, \tag{160}$$

[20] See J. Research NBS 45, 245 1950 RP2132. The variable "s" of Milne is changed to x and his λ^2 to λ, to avoid conflicts with the notations of the present paper.

is exactly the factor that compensates for the transition from du/dx to $\Delta u/\Delta x$, if $u(x)$ is of the following form:

$$u(x)=C_i \sin (\nu_i x+\theta_i), \tag{161}$$

where the constants C_i and θ_i are arbitrary.

Now it so happens that for large frequencies ν_i the first term of the operator (158) strongly overshadows the other terms. The differential equation of the eigenvalue problem for large ν_i thus becomes asymptotically:

$$\frac{d^2u}{dx^2}+\nu_i^2 u_i=0, \tag{162}$$

the solution of which is given by (161). This asymptotic behavior of the solution for large frequencies makes it possible to counteract the damaging influence of the error caused by the initial transition to the difference equation. The correction is implicitly included in Milne's solution, while the results of the matrix iteration scheme can be corrected by solving eq 159 for the ν_i.[21]

XIII. More-Dimensional Problems

The present investigation was devoted to differential and integral operators that belonged to a definite finite range of the variable x. This variable covered a one-dimensional manifold of points. However, in many problems of physics and engineering the domain of the independent variable is more than one-dimensional. A few general remarks may be in order as to the possibility of extending the principles and methods of the present investigation to manifolds of higher dimensions.

Although the general theory of integral equations reveals that the fundamental properties of an integral equation are essentially independent of the dimensionality of the variable x, yet from the practical viewpoint the eigenvalue problem of higher dimensional manifolds does lead to difficulties that are not encountered in manifolds of one single dimension. The basic difference is that an essentially more-dimensional manifold of eigenvalues is projected on a one-dimensional manifold, thus causing a strong overlapping of basically different vibrational modes. A good example is provided by the vibrational modes of a rectangular membrane. The eigenvalues are here given by the equation

$$\lambda=\alpha_1^2 m_1^2+\alpha_2^2 m_2^2,$$

where m_1 and m_2 are two independent integers, while α_1 and α_2 are two constants determined by the length and width of the membrane.

As another illustration, consider the bewildering variety of spectral terms that can be found within a very narrow band of frequencies, if the vibrational modes of an atom or a molecule are studied. To separate all these vibrational modes from each other poses a difficult problem that has no analogue in systems of 1 degree of freedom where the different vibrational states usually belong to well separated frequencies.

It is practically impossible to expect that one single trial function shall be sufficient for the separation of all these vibrational states. Nor does such an expectation correspond to the actual physical situation. The tremendous variety of atomic states is not excited by one single exciting function but by a rapid succession of an infinite variety of exciting functions, distributed according to some statistical probability laws. To imitate this situation mathematically means that we have to operate with a great variety of trial functions before we can hope to entangle the very dense family of vibrational states associated with a more than one-dimensional manifold.

In this connection it seems appropriate to say a word about the physical significance of the "trial function" $\phi_0(x)$ that we have employed for the generation of an entire system of eigenfunctions. At first sight this trial function may appear as a purely mathematical quantity that has no analogue in the physical world. The homogeneous integral equation that defines the eigenvalues, and the eigenfunctions, of a given integral operator, does not come physically into evidence since in the domain of physical reality there is always a "driving force" that provides the right side of the integral equation; it is thus the inhomogeneous and not the homogeneous equation that has direct physical significance.

If we carefully analyze the method of successive approximations by which the eigenvalues and the eigenfunctions of a given integral operator were obtained, we cannot fail to observe that we have

[21] This experience is valuable, since in many eigenvalue problems similar conditions hold; the eigenfunctions of large order can often be asymptotically estimated, in which case the error of the Δ-process may be effectively corrected. For example the values μ_i found in section IX for the lateral vibrations of an inhomogeneous bar may be corrected as follows:

μ_i uncorrected: 2256.944, 48.2038, 5.3563, 1.5830, 0.59, (0.25)

μ_i corrected: 2258.924, 48.4977, 5.4577, 1.6407, 0.62, (0.28)

basically operated with the *inhomogeneous* equation (103) and our trial function $\phi_0(x)$ serves merely as the "exciting function" or "driving force." Indeed, the solution (107) for the eigenfunctions is nothing but a special case of the general solution (102), but applied to such values of the parameter λ, which make the denominator zero. This means that we artificially generate the state of "resonance", which singles out one definite eigenvalue λ_i and its associated eigenfunction $\phi_i(x)$.

From this point of view we can say that, while the separation of all the eigenfunctions of a multi-dimensional operator might be a practically insuperable task—except if the technique of "separation" is applicable, which reduces the more-dimensional problem to a succession of one-dimensional problems—yet it might not be too difficult to obtain the solution of a given more-dimensional integral equation if the right side (i. e., physically, the "driving force") is given as a sufficiently smooth function that does not contain a too-large variety of eigenfunctions. Then the convergence of the method may still suffice for a solution that gives the output function with a practically satisfactory accuracy. This is the situation in many antenna and wave-guide problems that are actually input-output problems, rather than strict resonance problems. In other words, what we want to get is a certain mixture of weighted eigenfunctions, which appear physically together, on account of the exciting mechanism, while the isolation of each eigenfunction for itself is not demanded. Problems of this type are much more amenable to a solution than problems that demand a strict separation of the infinite variety of eigenfunctions associated with a multi-dimensional differential or integral operator. To show the applicability of the method to problems of this nature will be the task of a future investigation.

XIV. Summary

The present investigation establishes a systematic procedure for the evaluation of the latent roots and principal axes of a matrix, without constant reductions of the order of the matrix. A systematic algorithm (called "progressive algorithm") is developed, which obtains the linear identity between the iterated vectors in successive steps by means of recursions. The accuracy of the obtained relation increases constantly, until in the end full accuracy is obtained.

This procedure is then modified to the method of "minimized iterations", in order to avoid the accumulation of rounding errors. Great accuracy is thus obtainable even in the case of matrices that exhibit a large dispersion of the eigenvalues. Moreover, the good convergence of the method in the case of large dispersion makes it possible to operate with a small number of iterations, obtaining m successive eigenvalues and principal axes by only $m+1$ iterations.

These results are independent of the order of the matrix and can thus be immediately applied to the realm of differential and integral operators. This results in a well-convergent approximation method by which the solution of an integral equation of the Fredholm type is obtained by successive iterations. The same procedure obtains the eigenvalues and eigensolutions of the given integral operator, if these eigensolutions exist.

In the case of differential operators, the too-stringent demands of the least square method may be relaxed. The approximate linear identity between the iterated functions may be established by interpolation, thus dispensing with the evaluation of definite integrals. Moreover, the iterations may be carried out with the given differential operator itself, instead of reverting to the Green's function, which is frequently not available in closed form. The entire procedure is then free of integrations and requires only the solution of linear equations.

The present investigation contains the results of years of research in the fields of network analysis, flutter problems, vibration of antennas, solution of systems of linear equations, encountered by the author in his consulting and research work for the Boeing Airplane Co., Seattle, Wash. The final conclusions were reached since the author's stay with the Institute for Numerical Analysis, of the National Bureau of Standards. The author expresses his heartfelt thanks to C. K. Stedman, head of the Physical Research Unit of the Boeing Airplane Co. and to J. H. Curtiss, Acting Director of the Institute for Numerical Analysis, for the generous support of his scientific endeavors.

XV. References

[1] H. Hotelling, Psychometrika **1**, 27 to 35 (1936).

[2] A. C. Aitken, Proc. Roy. Soc. Edinburgh **57**, 269 to 304 (1937).

[3] W. Wayland, Quarterly of Appl. Math. **2**, 277 to 306 (1945).

[4] W. U. Kincaid, Quarterly of Appl. Math. **5**, 320 to 345 (1947).

[5] I. Fredholm, Acta Math. **27**, 365 to 390 (1903).

[6] Whittaker-Watson, Modern analysis, 4th ed., p. 221 (Cambridge Univ. Press, 1935).

[7] Whittaker-Watson, Modern analysis, 4th ed., p. 228 (Cambridge Univ. Press, 1935); Courant-Hilbert, Methoden d. math. Physik. **I**, 116 (J. Springer, Berlin, 1931).

[8] Courant-Hilbert, Methoden d. math. Physik. I. 76 (J. Springer, Berlin, 1931).

[9] Lord Rayleigh, The theory of sound **1**, 272 to 278 (reprinted; Dover Publications, New York, N. Y., 1945).

Los Angeles, September, 1949.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2134

Measurement of the Internal Diameters of Metallic Capillary Tubes

By Charles T. Collett, John C. Hughes, and Francis C. Morey

As a part of a cooperative research project of the National Bureau of Standards and the National Advisory Committee for Aeronautics on the viscosity of gases at elevated temperatures and pressures, it was found desirable to study the uniformity of the internal diameters of some metallic capillary tubes. This was accomplished by means of a small thread of mercury, using X-ray technique, following the method of Fisher. Eight tubes out of a group of twelve were selected as suitable for use in an absolute viscosimeter, and flow constants were computed. The tubes studied were about 15 feet long, with an internal diameter of about 0.015 inch. Maximum deviation from the average diameter in most cases was about 1.5 percent or less.

I. Introduction

A current cooperative research project of the National Bureau of Standards with the National Advisory Committee for Aeronautics involves the determination of the viscosity of gases at elevated temperatures and pressures, with particular reference to exhaust gases.

The method selected for the measurements is the classical capillary flow method, and in this connection the need for the determination of the internal diameters of some metallic capillary tubes arose. For use in absolute determinations, it is necessary to know not only the average internal diameter of the tube, but also the effect of irregularities of the diameters upon the flow characteristics.

Fisher[1] solved this problem with glass capillaries by measuring the length of a small thread of mercury placed in the tube and moved along it by air pressure. Variations in the cross-sectional area are detected as changes in length of the mercury thread. This method is used in selecting uniform capillary tubes for high-grade thermometers. Since the mercury thread was very small, Fisher did not attempt to determine its volume by weighing, but calculated it from the weight of a much longer thread by a method that will be considered in detail later.

The viscosimeter capillaries selected are metallic tubes of internal diameters about 0.015 in., wall thicknesses 0.005 to 0.014 in., and about 15 ft long. Manufacturers of such tubing were unable to supply any data on variations of diameter of commercial tubing except those based on experience in measurements made on the ends of the tubes. All felt confident that the variations would be of the order of ± 0.0005 in. or less. It was therefore decided to try to apply Fisher's method to the problem by making X-ray photographs of mercury threads in the tubing.

II. Description of Apparatus and Method

1. The short-thread measurements

The apparatus used in the X-ray inspection of the capillary tubes is shown in figures 1 and 2. The X-ray head, *A*, was suspended from a support so as to provide a beam directed vertically downward, through a slot in a lead shield, *B*. The capillary tube was carried on a frame, *C*, arranged to slide under the shield. A clamp, *D*, and blocks, *E*, were provided to facilitate advancing the tube 1 in. between exposures.

The capillary tube may be seen in figure 2 at *F*. It was held under slight tension in the frame and taped with cellulose tape to cross supports at intervals of about 3 ft. A plate-holder, *G*, could be clamped to the lower side of the frame, holding the glass photographic plate, *H*, enclosed in black paper envelope, tightly against the frame and in contact with the capillary tube.

The mechanism used to advance the mercury thread is seen in figure 2. A plastic cylinder, *I*,

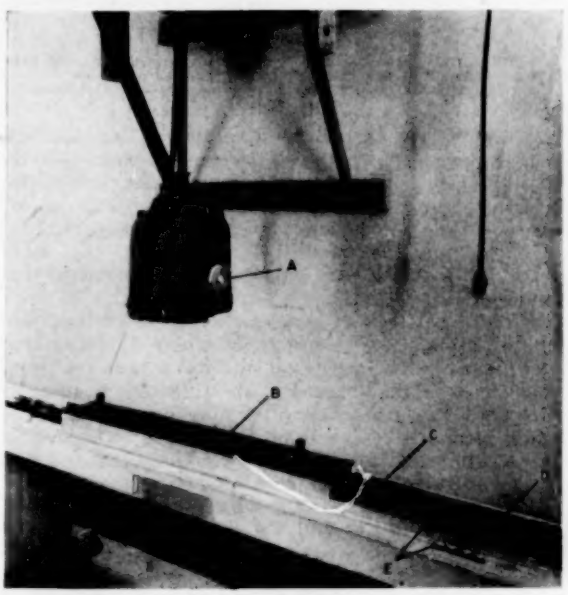

FIGURE 1. *General view of apparatus.*

A, X-ray source; *B*, lead shield with exposure aperture; *C*, rack for holding capillary tube; *D*, adjustable stop for positioning rack; *E*, removable blocks for positioning rack.

[1] W. J. Fisher, Phys. Rev. **28**, 73 (1909).

was attached to the capillary tube. A plunger operated by a micrometer screw, *J*, entered the cylinder through a packing sleeve. The cylinder was filled with a petroleum solvent oil, which entered the capillary and caused the mercury thread to move along the tube. A cushion of air about 4 in. long was always maintained between the thread and the oil. The plunger diameter was selected so as to cause the thread to advance 1 in. for each five divisions on the micrometer screw.

It was not found possible to rely solely upon the settings of the micrometer screw in advancing the mercury thread, as any slight drift due to irregularities of bore, temperature changes, or other factors

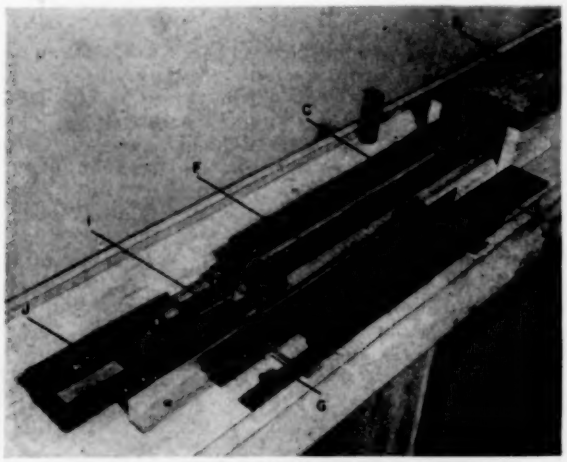

FIGURE 2. *General view of apparatus.*

B, Lead shield; *C*, rack for holding capillary tube; *F*, capillary tube; *G*, plate holder with pressure springs; *H*, photographic plate in paper envelope; *I*, liquid cylinder for positioning mercury slug; *J*, micrometer screw for advancing plunger.

might cause the thread to travel out of the exposure field. For this reason, strips of film were placed under the supporting table, about 3 in. below the capillary and exposed simultaneously with the plate. These check films were developed in a rapid developer and examined before making the next setting. Every fourth exposure was checked in this fashion.

The photographs were made on spectrographic plates, 2 by 18 in., on each of which 17 exposures were made. An overlap of one frame was provided on successive plates, and in these frames a series of identification notches cut into the supporting framework provided a means of automatically recording the position and for carrying over the measurements from one plate to the next. All measurements were referred to a small marker, which was hard-soldered to the tube near one end.

The apparatus was located in a subbasement corridor, in which temperature variations were found to be less than 0.5° C during the 6-hr period required for completing the 190 *X*-ray photographs of each tube. This was found to be sufficiently uniform to make temperature corrections unnecessary.

Thread lengths ranged from approximately 0.25 to 0.40 in. Measurements of the lengths of the thread images were made with a projection microscope at a magnification of about 50 diameters. In this apparatus, the plate being measured is placed upon a mechanical stage, displacement of which can be measured with a micrometer screw reading to 0.0001 in. It was found that repeated measurements of the same photograph differed from the average value by about ± 0.0002 in. The possibility of shrinkage of the emulsion during processing and handling was eliminated by using glass plates and avoiding use of the portions near the edges. During the first part of the work, a ball of known diameter was photographed in each frame and its image measured. When no shrinkage was observed in many photographs, its use was discontinued.

Cleaning agents were passed through the tubing with the aid of the house vacuum, and drying was accomplished by a current of air with moderate heating obtained by passing an electric current through the tube. The tube was degreased with trichloroethylene, dried; treated with concentrated nitric acid with heating nearly to boiling; rinsed with distilled water and thoroughly dried.

2. The long-thread measurements

As previously mentioned, it is not practical to determine the volume of the short mercury threads by direct weighing, as the mass is so small in proportion to a suitable container. It was therefore necessary to calculate this volume from the weight of a long mercury thread. The filling was accomplished by a vacuum method that was found necessary to eliminate air pockets. The tube was attached to a rotary oil-sealed vacuum pump. The opposite end was welded shut in a helium atmosphere with a tungsten electrode. The tube was then alternately evacuated and flushed with helium, while being heated electrically. This thoroughly dried the tube.

The welded closure was then removed and the open end of the tube inserted in a mercury reservoir made from a thin-walled glass tube 2 or 3 mm in diameter and about 15 cm long, which had been partially filled with mercury by vacuum distillation and had been weighed. The reservoir had a small bulb at the lower end so that air that might be carried down into the mercury by capillary action between the metal tube and the glass wall of the reservoir could not reach the end of the tube when fully immersed.

The tube extended vertically from the reservoir for a distance of about 3 ft and then horizontally to the pump. When the pump was started, mercury entered the tube and partially filled the vertical section, acting as a barometer. Evacuation was continued for 10 to 15 min to insure thorough removal of the air above the mercury. The tube was then welded closed at the pump end by using the helium atmosphere and tungsten electrode. Next, the mercury reservoir was slowly raised, permitting the mercury to flow in and fill the horizontal portion of the tube.

When the entire tube was horizontal and mercury had ceased entering the capillary tube, the tube was withdrawn from the mercury, but the end was kept within the reservoir. By gently warming the capillary tube, several drops of mercury were expelled and returned to the reservoir. Upon cooling, this left an air space several inches long near the end of the tube. The tube was then closed by welding. The mercury reservoir was again weighed. The weight of mercury entering the tube was about 5 to 7 g, and the initial weight of the reservoir approximately 15 g, permitting a good measurement of the weight of mercury entering the tube.

After filling, the length of the thread was determined from X-ray photographs, which were made without moving the capillary between exposures, so as to minimize the opportunity for movement of the thread inside the tube between exposures. After the end sections had been photographed, the entire tube was X-rayed and the photographs examined under magnification of 50 diameters to detect possible voids in the thread.

III. Method of Calculation

The following method of calculation (except for corrections for height of meniscus) is essentially that used by Fisher. The equation for laminar flow of a gas through a capillary tube, omitting corrections for slip, end effects, and expansion is

$$F=\frac{\pi r^4}{16l}\frac{(p_1^2-p_2^2)}{p\eta} \tag{1}$$

where

F = volume of gas discharged per unit time, measured at pressure p
r = radius of tube
l = length of tube
p_1 = entrance pressure
p_2 = exit pressure
p = average of p_1 and p_2
η = viscosity of the gas.

This may be rewritten

$$(p_1^2-p_2^2)=\frac{16pF\eta l}{\pi r^4}. \tag{2}$$

The tube may be considered to be made up of elements in series, having slightly varying radii. For each elemental length, dl,

$$d(p^2)=\frac{16pF\eta}{\pi}\frac{dl}{r^4}, \tag{3}$$

whence

$$\int d(p^2)=p_1^2-p_2^2=\frac{16pF\eta}{\pi}\int_{l_1}^{l_2}\frac{dl}{r^4}, \tag{4}$$

where d is the diameter of the capillary tube (inside).

Evaluation of the integral in the right-hand member is accomplished by measurement of the mercury thread. Neglecting, temporarily, the corrections due to the curvature of the ends, and assuming that we are dealing with a thread having flat ends, we let

λ = length of the thread
Q = cross-sectional area of the tube
v = volume of the thread
$Q=\pi r^2=v/\lambda$
$\frac{l}{r^4}=\frac{\pi^2\lambda^2}{v^2}$.

Therefore,

$$\int_{l_1}^{l_2}\frac{dl}{r^4}=\frac{\pi^2}{v^2}\int_{l_1}^{l_2}\lambda^2 dl. \tag{5}$$

If the observed values of λ^2 are plotted as a function of the position of the center of the thread from an arbitrary origin on the tube, the value of $\int_{l_1}^{l_2}\lambda^2 dl$ may be found graphically. The evaluation of v by direct weighing is difficult, because the sample is very small; in this case, averaging about 0.01 g. The volume may be calculated from the weight of a long mercury thread by the following means.

The volume of an elemental section of the tube is related to the cross-sectional area of the tube by the equation

$$dV=Qdl=\frac{v}{\lambda}dl.$$

The volume, V, of any portion of the tube can be found by integrating between appropriate limits.

$$V=v\int_{l_3}^{l_4}\frac{l}{\lambda}dl. \tag{6}$$

Solving this equation for the volume of the thread,

$$v=\frac{V}{\int_{l_3}^{l_4}\frac{l}{\lambda}dl}. \tag{7}$$

The volume, V, may be measured by means of a relatively long mercury thread, whose weight is great enough to be determined accurately by weighing. The integral is evaluated graphically. By substituting eq 7 in eq 5

$$\int_{l_1}^{l_2}\frac{dl}{r^4}=\frac{\pi^2}{V^2}\int_{l_1}^{l_2}\lambda^2 dl\left[\int_{l_3}^{l_4}\frac{l}{\lambda}dl\right]^2, \tag{8}$$

Combining eq 4 and 8 and solving for η gives

$$\eta=\frac{p_1^2-p_2^2}{pF}\frac{V^2}{16\pi\int_{l_1}^{l_2}\lambda^2 dl\left[\int_{l_3}^{l_4}\frac{l}{\lambda}dl\right]^2}. \tag{9}$$

The right-hand fraction thus becomes a calibration factor associated with the tube, corresponding to the constant $\pi r^4/16l$ in eq 1.

$$K=\frac{V^2}{16\pi\int_{l_1}^{l_2}\lambda^2 dl\left[\int_{l_3}^{l_4}\frac{l}{\lambda}dl\right]^2}, \tag{10}$$

where K is the flow constant.

Corrections for curvature of meniscus. The photographs of the mercury threads, when examined under a magnification of 50 diameters, were found to have rounded menisci, which in most cases were nearly hemispherical. Several examples are shown in figure 3, ranging from nearly perfect to badly distorted. As accurate measurement of the cylindrical portion of a thread would be difficult to make, it was decided to measure the over-all length and to reduce this observation to the length of a thread of the same total volume having flat ends. This involves measurement of the heights of the curved menisci and the difficulty of observing the point of contact of the mercury surface with the tube wall, but since the corrections are small, this measurement need not be made with great precision. Furthermore, it was found possible to draw a circle on the screen of the projection microscope, and to use this as a test for curvature of the menisci. Those which conformed closely to the circle could be assumed to be hemispherical and the meniscus height, equal to the average radius of the tube, used in computing the correction. About 80 percent of the photographs were so treated. In the remaining cases, the heights were measured and corrections computed by the following method.

In figure 4, h represents the observed meniscus height and r the tube radius. The shaded portion represents a cylindrical element having the same volume as the spherical segment. The volume, v, of a segment of a sphere of height h and radius of base r is

$$v=\frac{1}{6}\pi h(h^2+3r^2). \tag{11}$$

This volume is equal to that of the cylindrical element, shaded in figure 4. As we wish to determine the change in length that the thread would undergo if it were altered from a curved to a flat end, we may denote the change of length by S, giving

$$\pi r^2(h-S)=\frac{1}{6}\pi h(h^2+3r^2). \tag{12}$$

Solving for S gives

$$S=\frac{h}{6}\left(3-\frac{h^2}{r^2}\right). \tag{13}$$

Corrections varied from about 0.0034 in. for a well-rounded meniscus such as shown at A, figure 3, to about 0.0012 in. for a badly flattened one. The correction for case B, figure 3, was 0.0014 in.

Correction for target focus distance. Because of the finite distance of the X-ray source from the photographic plate, the image of the thread is slightly enlarged. The amount of this correction was minimized by keeping the tube in close contact with the paper envelope enclosing the photographic plate, and by having the distance from the plate to the X-ray target as great as permissible with reasonable exposure time. The greatest magnification factor encountered was 1.0016, giving a maximum correction of about 0.0005 in.

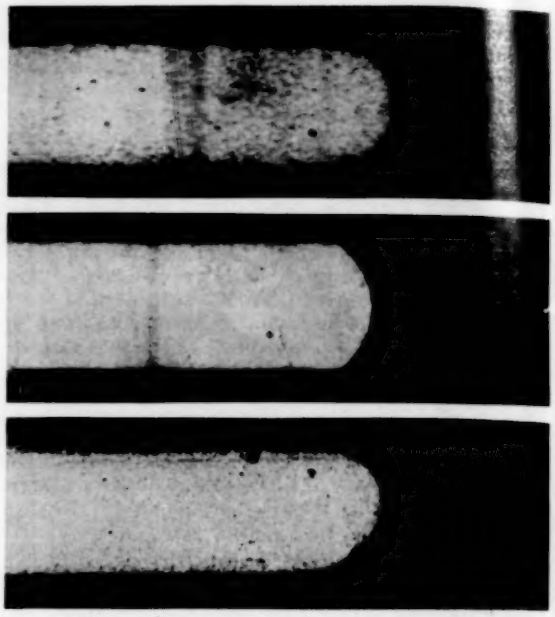

FIGURE 3. *Examples of typical photographs.*

A, Normal well-rounded meniscus; *B*, flattened meniscus; *C*, badly distorted meniscus.

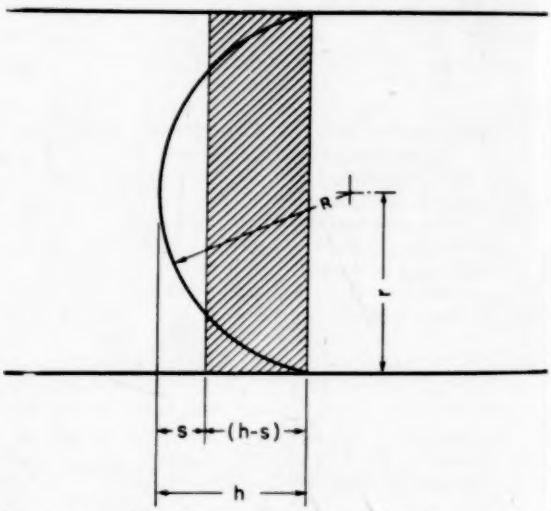

FIGURE 4. *Illustrating method of correcting for curvature of meniscus.*

IV. Results of Measurements

Figure 5 illustrates portions of typical graphs of λ^2 and $1/\lambda$ and the repeatability to be expected from consecutive trials with the same tube. Since the thread volume could not be duplicated exactly in the second trial, the plots are slightly displaced, but the general configuration correlates well.

Graphical integration was carried out in this case

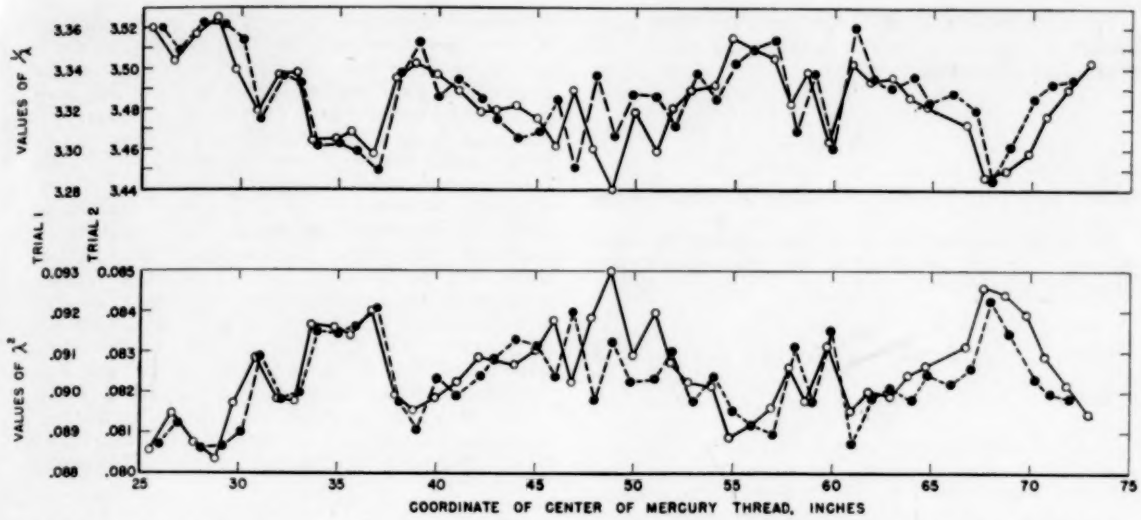

FIGURE 5. *Typical graphs of* $1/\lambda$ *and* λ^2 *used in graphical integration for calculating flow constants of capillary tubes.*

Two independent trials of the same tube illustrating degree of repeatability.

by a weighing method. The plots were prepared on strips of graph paper approximately 6 ft. long. The width of the strip depended upon the variations encountered in the function, but was about 6 to 12 in. After plotting, the paper was cut along the pencil line and along the limits of integration. Each piece of paper was compactly folded and was weighed after conditioning overnight in the case of the balance, to achieve humidity equilibrium. The ratio of the weight of the lower section of the paper to the combined weight of the two parts was computed. The combined area calculated from the dimensions of the strip was multiplied by the ratio of the weights to give the area beneath the curve.

This method of graphical integration is subject to errors due to nonuniformity of the paper, differences in the amount of inking of the ruling, and unequal soiling of the paper surface in handling. In this instance, the area integrated between the curve and the edge of the strip of paper was only a small percentage of the total area between the curve and the x-axis. In the integration of the λ^2 curves, the portion evaluated graphically was about 10 percent of the total area to the x-axis, whereas in the case of the $1/\lambda$ integrals, it amounted to only about 3 percent. As a further check, the areas for tube 2 were computed by the trapezoidal rule and by Simpson's rule, and the results of both methods of computation agreed with that obtained by the weighing method within ± 0.1 percent.

Values of the flow constant, K, are listed in column 2 of table 1. For comparison, column 3 lists values of $\pi r^4/16\ l$, using for r the average radius calculated from the weight of a long mercury thread.

Another method of illustrating the effect of irregularities in the tube is to compare the actual length of the tube with the length of an equivalent uniform tube whose radius is the same as that determined by the long mercury thread. These tube lengths are presented in columns 4 and 5.

TABLE 1. *Results of measurements of eight capillary tubes*

Tube number	Flow constant, K [a]	$\pi r^4/16l$	Length of tube	Length of equivalent tube of uniform bore [a]	Average inside diameter	Maximum plus departure from average diameter	Maximum minus departure from average diameter	Average departure from average diameter
1	2	3	4	5	6	7	8	9
			in.		*in.*	*Percent*	*Percent*	*Percent*
1 (trial 1)	3.375×10^{-12}	3.389×10^{-12}	160.00	160.66	0.01450	1.58	0.93	0.47
1 (trial 2)	3.380	3.389	160.00	160.43	.01450	1.46	1.11	.43
2	3.278	3.282	165.29	165.49	.01450	.98	1.28	.29
3	2.954	2.961	168.00	168.39	.01490	1.13	.45	.21
8	3.660	3.658	172.69	172.60	.01506	1.49	1.10	.35
9	3.782	3.793	174.00	174.50	.01523	2.45	1.06	.44
10	3.646	3.657	174.00	174.52	.01509	1.24	1.08	.36
11	3.609	3.609	174.00	174.00	.01504	1.08	1.66	.39
12	3.751	3.753	172.00	172.09	.01514	1.20	1.40	.34

[a] The values presented in these columns are subject to an uncertainty of approximately ± 0.2 percent.

For any position of the short mercury thread, the diameter is inversely proportional to the square root of the thread length.

$$\frac{d_1}{d_0}=\sqrt{\frac{\lambda_0}{\lambda_1}}, \tag{14}$$

where d_0 is the average diameter of the capillary tube.

We may compute the percentage deviation of the diameter at any point from the average diameter from the equation

$$D=100\left(\sqrt{\frac{\lambda_0}{\lambda}}-1\right), \tag{15}$$

in which

D=the percentage departure.

λ=the observed thread length.

λ_0=the average of the observed thread lengths.

Columns 7 and 8 of table 1 list the maximum positive and negative departures from the average diameter, and column 9 lists the average departure of the individual determinations from the average diameter.

The tubes studied were not all made by the same manufacturing process. The starting material in each case was tubing having an outside diameter of approximately 1¼ in. Tubes 1 to 3 were drawn over steel mandrels through dies that reduced the diameter by about one-quarter at each step, until the inside diameter was about 0.09 in. Thereafter the reductions were accomplished without mandrels.

A second set of tubes (not listed in table 1) were obtained from a manufacturer who drew the tubes to the finish diameter over mandrels. It was found that the rolls used in loosening the tubing from the mandrels were not perfectly centered and therefore produced a periodic variation in the diameters, which rendered the tubes unuseable. The variations amounted to as much as ±5 percent. Attempts were made to eliminate these differences by successive drawing through external dies, with slight diameter reduction. After four stages, the variations were somewhat reduced, but the periodic character persisted and the tubes were abandoned. Tubes 8 to 12, with heavier walls, were then supplied by the same manufacturer. In a lot of six tubes, only one was found to have unsatisfactory characteristics.

V. Summary

It has been found practical to apply the method of Fisher to the determination of flow constants for metallic capillary tubing and to study the uniformity of such tubing by X-ray photography of a mercury thread within the tube. Eight pieces of tubing were selected from a group of twelve, having a maximum deviation from the average diameter of 2.5 percent. Flow constants for these eight tubes were determined.

WASHINGTON, April 18, 1950.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2135

Mass Spectra of Some Simple Isotopic Molecules

By Vernon H. Dibeler, Fred L. Mohler, Edmund J. Wells, Jr., and Robert M. Reese

The mass spectra of H_2, D_2, T_2, HD, and HT have been measured with a Consolidated mass spectrometer with 50-volts ionizing voltage and with a constant magnetic field. The following ratios of atom ions to molecule ions were observed: $H^+/H_2^+=0.0201$, $D^+/D_2^+=0.0096$, $T^+/T_2^+=0.0061$, $H^+/HD^+=0.0096$, $D^+/HD^+=0.0097$, $H^+/HT^+=0.0085$, and $T^+/HT^+=0.0068$. The ratios for T_2 and HT were derived from spectra of two samples containing 91 and 37 atomic percent of T in equilibrium with H. Theoretical values for the ratios were computed for ionization involving transition into the Σ_g state of the molecule ion. These values are lower than the observed ratios, because transitions to higher repulsive states contribute to the observed values at 50-volts ionizing voltage.

Mass spectra of CO and CO_2 containing about 54 atom percent C^{13} were measured under standard operating conditions. For CO the ratio $C^{12+}/C^{12}O^+=0.0540$ and $C^{13+}/C^{13}O^+=0.0514$. For CO_2, the ratios are $C^{12}O^+/C^{12}O_2^+=0.0868$, $C^{13}O^+/C^{13}O_2^+=0.0824$, $C^{12+}/C^{12}O^+=0.0564$, and $C^{13+}/C^{13}O_2^+=0.0542$.

A theory of the isotope effect in CO involving simplifying assumptions indicates that the relative probability of producing C^{12+} is 3.5 percent greater than the probability of producing C^{13+}, as compared with an observed difference of 5 percent.

I. Isotopic Hydrogen Molecules

1. Introduction

Mass spectra of H_2 and D_2 have been investigated frequently [1],[1] and a recent paper by Friedel and Sharkey [2] gives the spectrum of HD. The present research was undertaken because of the opportunity to obtain some very pure HD and a sample of tritium of relatively good purity. Hydrogenic spectra are of interest because it is possible to compute theoretically from the Franck-Condon principle the relative probabilities of producing atom ions from the differ-

[1] Figures in brackets indicate the literature references at the end of this paper.

ent isotopic molecules, and a comparison of theory and experiment gives information on the mechanism of dissociation. Stevenson [3] has carried out the theoretical computations for H_2 and D_2. Data on these molecules are also of practical importance in the analysis of isotopic mixtures.

2. Experimental Procedure

A 180° Consolidated mass spectrometer was used, operated with a constant magnetic field requiring 3,800 v to focus mass 1 and using an ionizing voltage of 50 v. Pumps were speeded up, and pressure in the ionization chamber was kept fairly low to maintain a high vacuum in the spectrometer. This is a necessary precaution [4] with hydrogen, but as a result the atom ion peaks are rather small for accurate measurement. There are no triatomic ions under these conditions.

Very pure D_2 and HD were prepared by fractionation at liquid hydrogen temperatures by Fookson, et al. [5] of this Bureau. HD was fractionated from crude HD prepared by the reaction of D_2O and lithium aluminum hydride [6]. The tritium was obtained from Argonne National Laboratory and was 91 percent of T and 9 percent of H. It was separated from He^3 16 days before the spectrum was measured. A small correction amounting to 0.16 percent was applied for He^3 on the basis of a half-life of 12 years for tritium and the measured sensitivity (current per unit pressure) of He^4.

Table 1 gives the experimental results under columns headed "Expt." Data for H_2 and D_2 are consistent with published results. Friedel and Sharkey [2] report for HD values of 0.82 percent for H^+ and 1.13 percent for D^+. The value for D^+ of HD in table 1 is an upper limit assuming H_2^+ is zero. The value of 0.97 percent is the lowest value obtained in several experiments, but H_2^+ may contribute an amount of the order of 0.1 percent to mass 2. In the tritium both T_2 and HT are present and contribute to the mass 3 peak. The separate contributions were evaluated by obtaining the spectrum of the original sample containing 9 percent of H, then adding hydrogen and equilibrating the mixture in contact with an incandescent filament to obtain a mixture of 67 percent of H and 33 percent of T and repeating the measurement. In the first spectrum T_2 was large and TH small, and in the second spectrum the reverse was true. From the two spectra, values are derived for both T^+/T_2^+ and T^+/TH^+.

Table 1 includes the sensitivity (current per unit pressure) for the molecule ions. They are nearly equal except for T_2, which is about 10 percent lower than the other values. This is not surprising, as the ion accelerating voltage is only a third as much for T_2 as for H_2.

3. Theory

The ratio of atom ions to molecule ions can be computed by means of the Franck-Condon principle for the case of a transition from the normal state of hydrogen to the stable Σ_g state of the molecule ion.

TABLE 1. *Mass spectra of hydrogen molecules*

m/e	H_2		HD		D_2		HT		T_2	
	Expt.	Theor.	Expt.	Theor.	Expt.	Theor.	Expt.	Theor.	Expt.	Theor.
1	2.01	1.9	0.96	0.7			0.85	0.6		
2	100		.97	.7	0.96	0.8				
3			100				.68	.6	0.61	0.3
4					100		100			
6									100	
Sensitivity [a]	20.0		21.2		19.5				17.8	

[a] Ion current per unit pressure in gas reservoir for the molecule ions in arbitrary units.

Dissociation into an atom ion occurs if a vertical line from the normal state intersects the potential energy curve of the molecule ion at a point above the dissociation limit of the molecule ion. That is, the fraction of molecules with atoms inside a critical internuclear distance r_c will dissociate on ionization and, for H_2, about 2 percent of the atoms fall inside r_c and give atom ions. The potential energy curves of the Franck-Condon diagram and r_c remain unchanged in isotopic molecules. However, the radial distribution of atoms in the normal state is narrower for the heavy isotopes. This materially reduces the small fraction of atoms that fall inside r_c. Stevenson [3] has published computations for H_2 and D_2, and we have extended the computations to the other isotopic molecules. The values included in table 1 are subject to appreciable computational error, as they depend critically on the shape of the potential energy curves.

Experimental values are not exactly comparable with the theoretical values, for atom ions can also be produced by transitions to higher repulsive states of the hydrogen ion; namely, the Σ_u state and the doubly ionized state. The most probable transition to a repulsive state requires about 10 v in excess of the threshold for appearance of atom ions and results in ions of high kinetic energy. These ions are not collected efficiently by the mass spectrometer, and consequently they make only a small contribution to the atom ion peak. Hagstrum and Tate [7] have studied mass spectra of diatomic molecules and velocity distribution of the atom ions and have interpreted the results in terms of Franck-Condon diagrams. Their results for H_2 show the magnitude of the contribution of the high velocity atom ions from repulsive states.

The contribution of atom ions from repulsive states accounts for the fact that experimental results for 50-v electrons are always somewhat greater than the theoretical values. In the dissociation of HD the H^+ ion will have twice the kinetic energy of D^+. In the mass spectrum obtained with a constant magnetic field the kinetic energy received from the electric field will also be twice as large for H^+, so both ions should be collected with equal efficiency. Similarly H^+ and T^+ of HT should be measured with equal efficiency. The experimental observation that T^+ is less abundant than H^+ is unexplained. In other respects the experimental data seem quite consistent with theoretical predictions.

A mixture of D_2 and tritium was equilibrated in contact with an incandescent filament, but the spectrum of the mixture containing six isotopic molecules in appreciable amounts was too complicated to permit derivation of the DT spectrum. Theoretical computations indicate that D^+ and T^+ will each be 0.35 percent of DT^+.

II. Mass Spectra of CO and CO_2

1. Introduction

One would expect to find isotope effects in the mass spectra of molecules other than hydrogen, but the isotope effects will be much smaller when there is a relatively small difference in the mass of the isotopes involved. There is ample evidence for such effects both in mass spectra and in thermal dissociation. Dibeler, Mohler, and Reese have reported such effects in the spectrum of nitrogen containing 65 percent of N^{15} [8]. Beeck, et al. [9] have reported that the dissociation probabilities of $C^{12}-C^{12}$ and $C^{12}-C^{13}$ bonds of molecule ions formed from propane-1-C^{13} differ by about 20 percent, whereas Stevenson, et al. [10] find that in the thermal cracking of this molecule there is an 8-percent greater frequency of rupture of $C^{12}-C^{12}$ bonds than of $C^{12}-C^{13}$ bonds. Thermal decomposition of oxalic acid to give CO, CO_2, and H_2O has been studied by Lindsay, McElcheran, and Thode [11], and on the basis of the distribution of C^{13} in CO and CO_2 they concluded that the dissociation of a $C^{12}-O^{16}$ bond is approximately 3.5 percent more frequent than the dissociation of a $C^{13}-O^{16}$ bond.

2. Experimental Procedure

We have measured the relative abundance of the ions containing C^{12} and C^{13} in the mass spectrum of carbon monoxide and of carbon dioxide with high concentration of C^{13}. Carbon dioxide was evolved in vacuum by the action of concentrated sulfuric acid on barium carbonate containing 54 atom percent of C^{13}. A portion of the product was almost completely reduced to carbon monoxide over hot zinc according to the method of Bernstein and Taylor [12]. Traces of carbon dioxide were removed in a liquid nitrogen trap.

Mass spectra were obtained by using conventional techniques. The energy of the bombarding electrons was nominally 50 v, and the temperature of the ionization chamber was about 245° C. Data reported here were measured at constant magnetic field and varying ion accelerating voltage. In the case of CO_2, however, spectra were also obtained at constant ion accelerating voltage by varying the magnetic field. These spectra differ slightly in the ratios CO^+/CO_2^+ and C^+/CO_2^+, but the relative amounts of isotopic ions were the same within the estimated experimental error.

3. Results

Table 2 gives the mass spectra of carbon monoxide and carbon dioxide containing 54 atom percent C^{13}. Corrections have been made for the contributions of oxygen isotopes to the isotopic peaks. Columns 1 and 2 give the mass to charge ratio and the ion species for carbon monoxide. Column 3 gives the abundance of each ion relative to the most abundant ion, $m/e=29$. The ratio $C^{12}/C^{12}O^{16}$ equals 0.0540, whereas the ratio $C^{13}/C^{13}O^{16}$ equals 0.0514. The ratio of these two values, 1.05, is a measure of the relative probability of dissociating a $C^{12}-O^{16}$ bond as compared with a $C^{13}-O^{16}$ bond in the CO molecule.

TABLE 2. *Mass spectra of* CO *and* CO_2 *containing approximately 54 atom percent of* C^{13}

[Correction has been made for oxygen isotopes]

Carbon monoxide			Carbon dioxide		
m/e	Positive ion	Relative abundance	m/e	Positive ion	Relative abundance
12	C^{12}	4.68	12	C^{13}	4.91
13	C^{13}	5.14	13	C^{13}	5.42
16	O^{16}	3.16	16	O^{16}	15.4
28	$C^{12}O^{16}$	86.7	28	$C^{13}O^{16}$	7.55
29	$C^{13}O^{16}$	100.0	29	$C^{13}O^{16}$	8.20
			44	$C^{12}O_2^{16}$	87.0
			45	$C^{13}O_2^{16}$	100.0

Columns 4, 5, and 6 of table 2 give the mass spectrum of carbon dioxide relative to the most abundant ion, $m/e=45$. For the dissociation of a single C—O bond, the ratio $C^{12}O^{16}/C^{12}O_2^{16}$ equals 0.0868, whereas the ratio $C^{13}O^{16}/C^{13}O_2^{16}$ equals 0.0824. Again, the ratio of these two values, 1.05, is a measure of the probability of dissociating a $C^{12}-O^{16}$ bond, as compared with a $C^{13}-O^{16}$ bond. For the dissociation of two C—O bonds, the ratio $C^{12}/C^{12}O_2^{16}$ equals 0.0564, whereas the ratio $C^{13}/C^{13}O_2^{16}$ is 0.0542. Rather surprisingly the ratio of these values, 1.04, indicates that the dissociation of two $C^{12}-O^{16}$ bonds is only 4 percent more frequent than the dissociation of two $C^{13}-O^{16}$ bonds. Measurements made on a sample of carbon dioxide containing 4 atom percent of C^{14} gave similar, though less accurate, results for the relative probability of dissociating $C^{12}-O^{16}$ bonds and $C^{14}-O^{16}$ bonds. The results on the preferential dissociation of a $C^{12}-O^{16}$ bond compared with a $C^{13}-O^{16}$ bond in the ionization process are similar to the results of Lindsay, McElcheran, and Thode [11] on thermal dissociation of oxalic acid.

4. Discussion

Similar effects of isotopes on the dissociation probabilities of other molecules have been observed in this laboratory. From analysis of the mass spectrum of nitrogen [8] containing 65 atom percent of N^{15} it was found that an $N^{14}-N^{14}$ bond dissociates about 11 percent more frequently than an $N^{15}-N^{15}$ bond. The relative probability of dissociating an $N^{14}-N^{15}$ bond was less accurately determined, but the value is intermediate between the values for N_2^{14} and N_2^{15}.

It is to be expected that isotope effects in CO and N_2 can be explained in terms of the Franck-Condon

principle, except for the difficulty that potential energy curves for the molecule ion states are not accurately known. D. P. Stevenson in an unpublished paper presented before the New Orleans meeting of mass spectrometer users in May 1950 pointed out that the isotope effect in N_2 can be estimated without detailed knowledge of the upper potential energy curve. The observation that N^+ is about 6.7 percent of N_2^{14+} indicates that in the radial distribution of atoms in the normal state, 6.7 percent of the atoms fall inside the critical internuclear distance for dissociation in the ionized state. From the known distribution of atoms in the normal state, the critical internuclear distance can be computed. In N_2^{15} the distribution of atoms will be reduced by an amount that will result in a 7-percent reduction of the fraction inside the critical internuclear distance. This is considered in substantial agreement with the observed difference of 11 percent.

This interpretation is supported by the research of Hagstrum and Tate [7] on the velocity distribution of atom ions as related to the Franck-Condon diagrams. The velocity distribution of N^+ ions indicates that dissociation comes primarily from a process of the type postulated by Stevenson. This is also true of C^+ from CO. On the basis that 5.4 percent of C atoms of $C^{12}O^{16}$ fall inside the critical internuclear distance, it is estimated that in $C^{13}O^{16}$ there will be a 3.5-percent reduction in the fraction inside the critical distance. The observed difference was 5 percent. Exact agreement is not expected, for the theory assumes that a single transition process gives rise to C^+. It is known that at least two transitions give C^+; one or more giving C^+ and O and one or more giving C^+ and O^-.

III. Conclusion

Isotope effects in the diatomic molecules, hydrogen, nitrogen, and C^+ of CO seem to be accounted for by the Franck-Condon principle and the fact that the radial distribution of atoms in the normal state depends on the mass of the isotopes. In these molecules a small fraction of the atoms in the normal state falls inside the critical internuclear distance for dissociation in the ionized state, and this condition leads to a relatively large isotope effect. The research of Hagstrum and Tate [6] on the ionization of diatomic molecules indicates that production of O^+ from O_2 and of N^+ from NO will be similar in this respect to N_2. However, for the transitions giving O^+ from CO and O^+ from NO, roughly half and three quarters respectively of the atoms fall inside the critical radius and the isotope effect would presumably be much less. One would not expect the relative amount of O^{16+} from the isotopic mixture of $C^{12}O^{16}$ and $C^{13}O^{16}$ of table 2 to be significantly different from normal CO, and this is experimentally observed.

Although similar effects may be expected in polyatomic molecules, it is not feasible to predict quantitatively the isotope effects in CO_2 or more complicated molecules. It may be coincidental that production of C^+ and CO^+ from CO_2 show almost the same isotope effect as C^+ from CO. The theory of diatomic molecules indicates that each dissociation process will have a characteristic isotope effect, and values of zero or even negative values are not excluded.[2]

IV. References

[1] N. Bauer and J. Y. Beach, J. Chem. Phys. **15,** 150 (1947).
[2] R. A. Friedel and A. G. Sharkey, Jr., J. Chem. Phys. **17,** 584 (1947).
[3] D. P. Stevenson, J. Chem. Phys. **15,** 409 (1947).
[4] R. E. Honig, J. Chem. Phys. **16,** 837 (1948).
[5] A. Fookson, P. Pomerantz, and E. H. Rich, Science, in press.
[6] I. Wender, R. A. Friedel, and M. Orchin, J. Am. Chem. Soc., **71,** 140 (1949).
[7] H. D. Hagstrum and J. T. Tate, Phys. Rev. **59,** 354 (1941).
[8] V. H. Dibeler, F. L. Mohler, and R. M. Reese, J. Chem. Phys. **18,** 156 (1950).
[9] O. Beeck, J. W. Otvos, D. P. Stevenson, and C. D. Wagner, J. Chem. Phys. **16,** 255 (1948).
[10] D. P. Stevenson, C. D. Wagner, O. Beeck, and J. W. Otvos, J. Chem. Phys. **16,** 993 (1948).
[11] J. G. Lindsay, D. E. McElcheran, and H. G. Thode, J. Chem. Phys. **17,** 589 (1949).
[12] R. B. Bernstein and T. I. Taylor, Science **106,** 498 (1949).
[13] O. Schaeffer and J. Hastings, Quarterly Progress Report, p. 31 (Brookhaven National Labratory, Mar. 31, 1950).

WASHINGTON, June 21, 1950.

[2] A progress report by O. Schaeffer and J. Hastings of the Brookhaven National Laboratory [13] on mass spectra of H_2, D_2, and T_2 has appeared since this paper was written. They report ratios of atom ions to molecule ions obtained from ionization by electrons of 30 electron volts, and values are considerably smaller than values reported in table 1 at 50 ev, as is to be expected. In other respects their results are comparable with ours.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2136

Response Characteristics of Temperature-Sensing Elements for Use in the Control of Jet Engines

By Andrew I. Dahl and Ernest F. Fiock

The rate at which a temperature-sensing element located in the gas stream of a jet engine responds to sudden changes in temperature is of great practical importance in the control and operation of such an engine. The factors that determine rate of response are discussed, and the significance of the characteristic time is emphasized. It is shown that laboratory determinations of characteristic time must be made under simulated engine conditions in which the rate of heat transfer by forced convection is the controlling factor. Apparatus used at the National Bureau of Standards for measuring characteristic times is described, and typical results are presented. The rate of response of a device in a jet engine varies greatly with engine speed and with the altitude of flight, so that satisfactory performance of a temperature-actuated control system can be expected only if the sensing element responds with sufficient rapidity under starting conditions and at the flight ceiling.

I. Introduction

The importance of accurate temperature measurements in the development and operation of gas turbines and jet engines is well recognized. Such measurements determine important operating characteristics of the power plant and permit the engine designer to single out for special attention those components wherein changes would be most effective in improving performance.

Of commensurate importance is the application of temperature-sensing devices in the control of jet engines. An acceptable control system must prevent the engine speed from increasing at an excessive rate and protect critically stressed parts from being overheated. Since a change in fuel-air ratio is normally accompanied by a simultaneous change in gas temperature, adequate protection of the power plant requires a control that will act very suddenly to limit the rate of change of gas temperature during acceleration and to reduce the gas temperature whenever it exceeds a safe value. If the control is temperature-actuated, the sensing element must therefore respond rapidly when subjected to a change in temperature.

Protecting against overheating of the part that is considered most likely to fail mechanically might be achieved by attaching a temperature-sensing element directly to such part. Usually, however, most highly stressed parts are rotating, so that it is extremely difficult to make satisfactory temperature measurements of such parts. Even if this installation problem were solved, the instrument so applied would respond but slowly to the stimulus (i. e., a change in gas temperature) that causes engine speed to change. Thus it is logical to base the control system upon the temperature of the high-velocity gas, since such a system can be made to protect against overheating, to prevent too rapid acceleration, and to establish the power output of the engine. Obviously, these essential controls can be achieved only if the temperature sensing element and the auxiliary parts constituting the control system can be made to respond very rapidly to changes in gas temperature.

In the course of projects dealing with temperature-sensing devices applicable in gas turbines and jet engines, sponsored at the National Bureau of Standards by the Air Matériel Command and by the Bureau of Ships, many misconceptions have come to light concerning the rate of response of such devices to sudden changes in temperature. These do not involve the basic theory of response, which has been treated clearly by Harper [1],[1] but indicate some confusion as to the significance of measured values and as to the proper application of such values.

This paper is concerned particularly with the rates of response of interest in the control of turbine-type engines, where the temperature-sensing element is normally immersed in flowing gas. Various factors that determine the rate of response under such circumstances are discussed, the relative importance of these factors is indicated, and laboratory equipment and methods used at this Bureau for determining response characteristics are described. Previous papers in this field [2, 3, 4, 5] deal with other applications, and therefore do not emphasize some of the factors that are specific to power plants, particularly those for aircraft.

II. General Considerations

An object immersed in a gas at temperature T_g, retained by walls at temperature T_w, eventually attains a steady temperature, T_2. At this steady state the object gains and loses heat at equal rates, and its actual temperature T_2 is seldom identical with either T_g or T_w. The object may be a temperature-sensitive element initially at any given temperature T_1, which is introduced suddenly within the aforementioned environment. The rate at which its temperature changes from T_1 to T_2 may now be measured and is found to be a function of its surface area (S), mass (M), heat capacity (C), and the rate of heat transfer with the surroundings. This may be expressed as follows:

$$dT/dt=(S/MC)f(k,T,T_g,T_w), \quad (1)$$

[1] Figures in brackets indicate the literature references at the end of this paper.

in which T is its temperature at time, t, after its environment is changed, and k is a generalized coefficient of heat transfer such that $f(k,T,T_g,T_w)$ describes completely the heat transfer to and from the element.

For every specific element, the heat-transfer function in turn governs the rate at which the element responds to sudden changes in environment. Although this function cannot be evaluated theoretically for the completely general case, present purposes are served by considering special cases in which the heat transfer is by one single method at a time, because, in most practical applications including that in engines, the rate of transfer by one method is so preponderant that the effects of all others are negligible by comparison.

III. Factors Influencing Response in Gas Streams

1. Heat Transfer by Convection

In the transfer of heat by fluid convection, motion of the molecules continuously brings new particles in contact with the surface of the object being heated or cooled. A distinction is made between natural convection, in which the motion of the medium is due to differences in density resulting from differences in temperature, and forced convection, in which the motion of the medium is produced by mechanical means. The transfer of heat by conduction through gas is included in the transfer by convection, so that it need not be treated separately. In any case, the rate of transfer of heat by conduction is small, compared with that by convection.

Consider, first, the case in which radiation can be neglected. This could be approximated closely if the object were provided with an efficient radiation shield. Under these conditions the wall temperature is without effect, and $T_2=T_g$. Thus eq 1 may be written

$$dT/dt=(S/MC)f(h_cT,T_2), \tag{2}$$

in which h_c is the coefficient of heat transfer by convection. Also, the rate of heat transfer by convection varies directly as the difference in temperature (Newton's law of cooling), so that $f(h_c,T,T_2)$ becomes $(T_2-T)f(h_c)$.

The value of h_c for forced convection is essentially independent of temperature, so that for this case eq 2 becomes

$$dT/dt=h_c(S/MC)(T_2-T). \tag{3}$$

It has been shown experimentally [6, 7] that, for natural convection, the coefficient may be expressed as

$$h_c=k_1(\Delta T)^{1/4}, \tag{4}$$

in which k_1 is a constant and ΔT is the temperature excess of the gas over the element heated. Thus, for this case eq 1 becomes

$$dT/dt=k_1(S/MC)(T_2-T)^{5/4}. \tag{5}$$

(a) Forced Convection

This is the case of real interest when considering the response of temperature-sensing elements immersed in the gas stream of jet engines, because the rate of heat transfer by forced convection far exceeds that by any other method in such applications.

With τ substituted arbitrarily for the quantity MC/Sh_c eq 3 integrates to

$$-\log_e T_2-\mathrm{T})=t/\tau+\text{ constant}, \tag{6}$$

and, since $T=T_1$ at $t=0$,

$$(T-T_1)/T_2-T_1)=\Delta T/\Delta T_0=1-e^{-t/\tau}, \tag{7}$$

in which ΔT is the change in temperature that the element has experienced at time t, ΔT_0 is the total change to which it is subjected, and e is the base of Naperian logarithms.

From eq 7 it is apparent that the quantity τ must have the dimensions of time, and that at time $t=\tau$, the ratio $\Delta T/\Delta T_0=1-1/e=0.632$. Thus τ is the time required for the object to undergo 63.2 percent of the total change in temperature to which it is subjected instantaneously. The time τ, so defined, is generally referred to as the time constant or "characteristic time" of the element, and this expression will be adhered to throughout the remainder of this paper. A similar concept of the time constant appears frequently in the fields of mechanics and electricity.

The characteristic time of a temperature-sensing element has a number of interesting physical interpretations. Referring to a specific installation, both ΔT_0 and τ are independent of time, and eq 7 can be differentiated to give

$$d(\Delta)Tdt=(\Delta T_0/\tau)e^{-t/\tau}. \tag{8}$$

Equation 7 may also be written

$$T_2-T=\Delta T_0e^{-t/\tau}. \tag{9}$$

Dividing eq 9 by eq 8 gives

$$(T_2-T)/[d(\Delta T)/dt]=\tau. \tag{10}$$

Equation 10 shows that the characteristic time may also be defined as the time that would be required for the object to change from temperature T at time t, to temperature T_2, if the rate of change $[d(\Delta T)/dt]$ of temperature prevailing at time, t, were held constant. This interpretation of the characteristic time is shown graphically in figure 1, in which the curve represents eq 7, plotted in terms of the dimensionless ratios $\Delta T/\Delta T_0$ and t/τ. It will be noted that if a tangent is drawn to the curve at any point and extended to $\Delta T/\Delta T_0=1$ (or $T=T_2$), the horizontal component of such a tangent is equal to τ.

The quantity τ was introduced above as a substitute for the quotient MC/Sh_c. Among these quantities, M, C, and S are physical characteristics

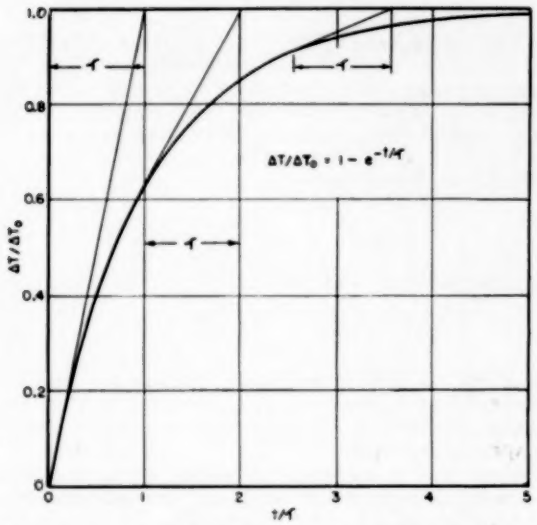

FIGURE 1. *Physical interpretation of characteristic time.*

Heat transfer by forced convection only.

of the object, and if the coefficient of heat transfer by forced convection (h_c) were independent of flow conditions, then the characteristic time of a particular object would be a single-valued property. Actually h_c varies in a complex manner, for which no satisfactory theoretical treatment has been evolved. However, comprehensive analyses of available data indicate that, for a given object, h_c is essentially dependent only on the mass rate of flow. Thus, the characteristic time of an object immersed in gas flowing at velocities attained in jet engines is essentially independent of the temperature and of the temperature difference to which it is subjected, and the same value of τ applies whether the object is being heated or cooled. For instruments used in flowing gases, a numerical value of characteristic time is significant only when the mass flow rate is specified, but no statement as to the magnitude or direction of the temperature change is required. For example, it is meaningful to say that a particular device has a characteristic time of 2 sec at a mass flow rate of five lb/sec ft^2.

Since the characteristic time of any object varies directly as its mass and inversely as its surface area, it is apparent that τ will be decreased by any process that increases the surface area proportionately greater than the mass. Consequently, any method of increasing the surface area without changing the mass will reduce τ, but such means for reducing τ are practicable only within the limits established by the ability of the element to withstand the stress to which it is subjected.

(b) Natural Convection

It is known that the rates of response of instruments designed for use in jet engines have been measured in apparatus in which they gained or lost heat primarily by natural convection and radiation. The following sections are included merely to show why values measured in this way cannot be significant, so far as applications in engines are concerned.

If the expression Sk_1/MC is represented by β, eq 5 may be written

$$dT/dt=\beta(T_2-T)^{5/4}, \tag{11}$$

which may be integrated to give

$$t=[4\beta/(T_2-T_1)^{1/4}]\{[(T_2-T_1)/(T_2-T)]^{1/4}-1\}. \tag{12}$$

It can be seen from eq 12 that the rate at which an object will respond to a sudden change in temperature, when heat is transferred solely by natural convection, is a function of the temperature interval over which the heating takes place. This is shown graphically in figure 2, in which $t/4\beta$ and $\Delta T/\Delta T_0$ are plotted for the single value $T_1=500°$ F and for three values of ΔT_0.

Since the response due to natural convection depends upon the particular temperature conditions encountered, whereas that due to forced convection does not, the characteristic times of the same device under these two sets of operating conditions cannot be related.

2. Heat Transfer by Radiation

Consider the special case in which heat transfer to or from a relatively small body, such as a thermocouple junction, takes place by radiation only, so that there is no transfer of heat by convection, Tg has no effect, and $T_2=T_w$. As an example, a junction might be moved suddenly from a cool to a hot location within a vacuum furnace. The transfer of heat in this case takes place in accordance with the Stefan-Boltzmann law, and eq 1 becomes

$$dT/dt=(S\sigma\epsilon/MC)(T_2^4-T^4)=\alpha(T_2^4-T^4), \tag{13}$$

in which σ is the Stefan-Boltzmann constant, ϵ is the surface emissivity of the junction, and α stands for the quotient $S\sigma\epsilon/MC$.

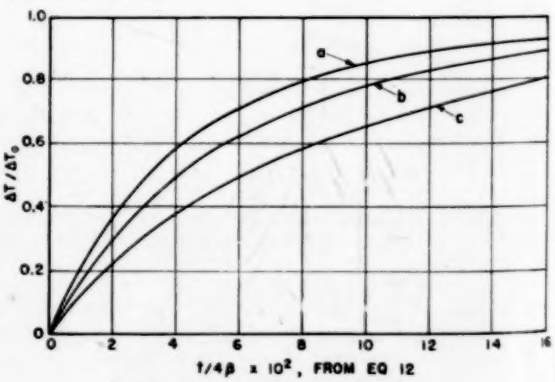

FIGURE 2. *Graph of equation 12.*

Heat transfer by natural convection only. Curve a, $\Delta T_0=1{,}500°$ F; curve b, $\Delta T_0=500°$ F; curve c, $\Delta T_0=100°$ F.

Integration of eq 13 gives

$$t=\frac{1}{4\alpha T_2^3}\left[\log_e \frac{(T_1-T_2)(T+T_2)}{(T_1+T_2)(T-T_2)}+2\tan^{-1}(T/T_2)-2\tan^{-1}(T_1/T_2)\right]. \quad (14)$$

Examination of this equation shows that the time-temperature relation, and hence also the rate of response, depends markedly on the following: (a) whether the junction is being heated or cooled; (b) the temperature level; and (c) the temperature interval.

Item (a) can be illustrated by considering that, in one case, a junction heated from 500° to 2,000° R, and in another it is cooled from 2,000° to 500° R. The quantity $4\alpha t$ is plotted against temperature in figure 3 for each of these cases. Although the initial

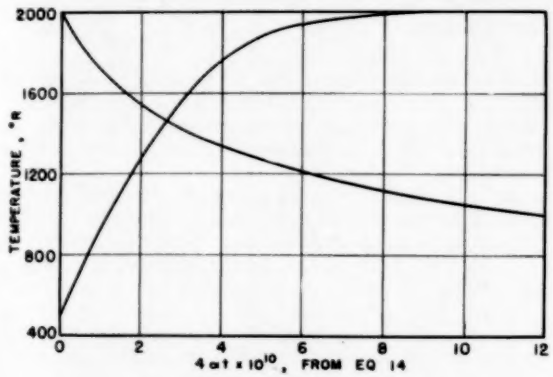

FIGURE 3. *Graph of equation 14.*
Heat transfer by radiation only.

rates of temperature change are the same in the two cases, the time required for the temperature to change by a given percentage of the temperature difference is less during heating than it is during cooling.

Thus when most of the heat is transferred by radiation, the rate of response of a given junction depends on different factors and has entirely different characteristics than when the primary heat transfer is by convection.

IV. Conditions of Heat Transfer in Jet Engines

The gas flow rates that prevail in a jet engine cover a wide range and depend primarily upon the engine operating conditions and the altitude of flight. However, even during starting, heat is transferred to a device immersed in the gas stream predominately by forced convection. Since the various solid parts that retain the gas stream are normally cooler than the gas because of heat losses, a sensing element in the gas stream normally attains a steady temperature below that of the gas itself, because of radiation and conduction from the sensing element. Nevertheless, when the temperature of the gas is changed suddenly, the rate of change in the temperature of the sensing element is determined almost entirely by the rate of heat transfer by forced convection, and the effects of radiation and conduction over the short interval of time involved are insignificant.

As the characteristic time of the instrument is a function of the operating conditions, as well as of the physical characteristics of the device itself, these conditions must be taken into account in devising laboratory equipment for measuring τ and in using measured values in the design of control systems for engines. The mass flow rate through a given engine, and hence also the characteristic time of an instrument immersed in its gas stream, varies with the density of the ambient air, engine rpm, and to a lesser extent with air speed. Thus the over-all rate of response of the control system must be sufficiently high that adequate protection is provided at the lowest mass flow rates, which occur during static starts and at the flight ceiling. A system that protects the engine at sea level and high-power outputs may be utterly inadequate at low-power outputs and at high altitudes.

V. Laboratory Apparatus for Measuring Characteristic Time

Equipment used at this Bureau for studying the performance of temperature-sensing devices applicable in jet engines is shown schematically in figure 4, and the apparatus used for producing sudden changes in the temperature of the test instruments is shown in figure 5. As indicated in figure 4, a compressor supplies air to a single Jumo 004 turbojet engine combustor equipped with its normal fuel injector and spark plug. The gas temperature and flow rate at the combustor outlet are subject to control through a valve at the inlet to the compressor and by adjusting the fuel pressure. Exhaust gases from the burner pass through two 90° turns, through a perforated plate, and then through about 10 ft of straight pipe before reaching the test section. The latter has three convenient hatches for mounting instruments in the gas stream, which is at essentially uniform temperature and velocity over the central half diameter of the test section. Due to the shape of the duct system, no instrument is exposed to direct radiation from the flame.

A bleed line containing a butterfly valve provides an auxiliary control of the flow in the test section, independent of the operating conditions of the burner. Pressures observed with a pitot-static tube, together with the known value of gas temperature, permit calculation of the mass flow rate in the test section.

Two such systems have been used to date, and it is planned to maintain both for future studies of a similar nature. The air for one is supplied by a blower with a capacity of 4,000 ft^3 of free air a minute and that for the other by centrifugal compressors having a combined capacity of 10,000

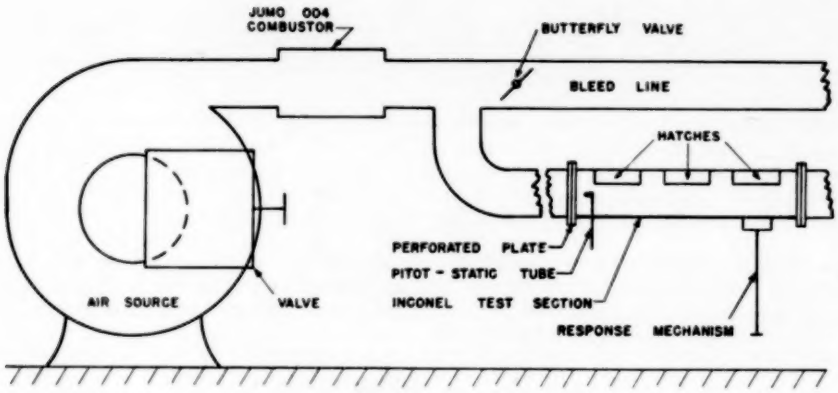

FIGURE 4. *Diagram of test systems.*

ft³/min. The systems have identical Inconel test sections, about 6 in. in diameter. The remainder of the former is of ordinary iron, which limits the operating temperature to about 1,600° F because of scaling. Most of the work on rates of response has been done in this system, in which the blower capacity and the drop in pressure through the burner and lines limit the mass flow rate to approximately 8 lb/sec ft². Currently a second blower is being added in series, and all iron pipe downstream of the burner is being replaced with Inconel, in order that the operating range may be increased.

All of the second system beyond the Jumo combustor is of Inconel, and an afterburner is installed at the burner outlet. This system can be operated at temperatures up to about 2,000° F, and at mass flow rates up to 15 lb/sec ft² in the 6-in. test section. A test section 3 in. in diameter is available for use in life tests and in research where velocities up to 1,800 ft/sec are of interest.

For measuring the rate of response, the temperature of the gas surrounding the test instrument must be changed suddenly. This cannot be accomplished with sufficient rapidity by changing the burner operating conditions. It can be done, without altering the steady flow of exhaust gas in the test section, with the apparatus shown in figure 5. An Inconel tube, held in position around the test instrument by a release plate, provides a flow channel for cold air. Upon removing the release plate, the Inconel tube is removed suddenly by a spring, thus exposing the instrument to the hot gas of the main stream almost instantaneously. During the downward movement of the tube, the supply of cold air is stopped automatically. In this way a test instrument at a known, moderate temperature (controlled by the air rate through the Inconel tube) can be exposed to exhaust gas at any chosen temperature and mass flow rate within the capability of the test system.

The performance of a variety of thermocouples, resistance thermometers, and thermistors has been studied. In each case the experimental result was recorded with a direct-inking oscillograph, in the form of an emf-time chart. The time scale can be had either by counting the strokes of the pen, which oscillates continuously at 60 cycles/sec, or from the known speed of the paper, which is moved by a synchronous motor.

An a-c amplifier was used with the oscillograph, so that it was necessary to use a "chopper" on the output of each thermocouple. With the resistance-type devices, both a-c and d-c bridges with choppers have been used with equal success. In any of these systems the amplitude of the recording pen is proportional to the amplified emf, so that the envelope of the record gives the impressed emf as a function of time. It is not necessary to know the amplification factor. The initial and final temperatures, if desired, can be measured more conveniently and more accurately with other indicating or recording instruments.

If the output of the sensing element varies approximately linearly with temperature, as is the case with thermocouples and resistance thermometers, the characteristic time can be read directly from the record. It is simply the time required for the amplitude to change from its initial value by 63.2 percent of the total change that it experiences.

A typical record of this type, together with a drawing of the envelope curve to an enlarged scale, is shown in figure 6. This particular record is for a No. 22 gage bare Chromel-Alumel thermocouple, the junction of which was formed by fusing the ends of the wires into a bead having a diameter approximately twice that of a single wire. Initial and final temperatures, measured with other instruments, were 300° and 1,200° F, respectively, and the mass flow rate was 4 lb/sec ft².

In the case of thermistors, for which the resistance decreases approximately exponentially with increasing temperature, a satisfactory arrangement is to use the test unit as one arm of a bridge, in which the oscillograph indicates balance. Having determined the resistance-temperature relation for the test instrument, and having chosen the experimental values of T_1 and T_2 to be used in the test, the resistance corresponding to the temperature $[T_1+0.632$

$(T_2 - T_1)$] is calculated and set up in a second arm of the bridge. The remainder of the test of a thermistor can then be conducted as described for thermocouples, but the oscillograph record now shows a maximum amplitude at the start of the test and a minimum amplitude at the characteristic time, as measured from the instant of exposure to hot gas.

A wide variety of instruments was studied, including sensing elements exposed directly to the gas stream, elements encased in metal and ceramic protection tubes, and elements imbedded in insulating materials such as quartz and beryllia. Characteristic times as low as 0.02 sec have been measured successfully, and there seems to be no upper limit imposed by the test system or recording equipment. For each instrument, from three to five records were taken at each of the flow rates 2, 4, and 6 lb/sec ft^2. No results have been obtained at flow rates above 6.8 lb/sec ft^2, primarily because the flow rates of greatest interest in engines are less than this.

Presentation of the results in detail would be useless without a complete description of each instrument, which is considered neither practicable nor

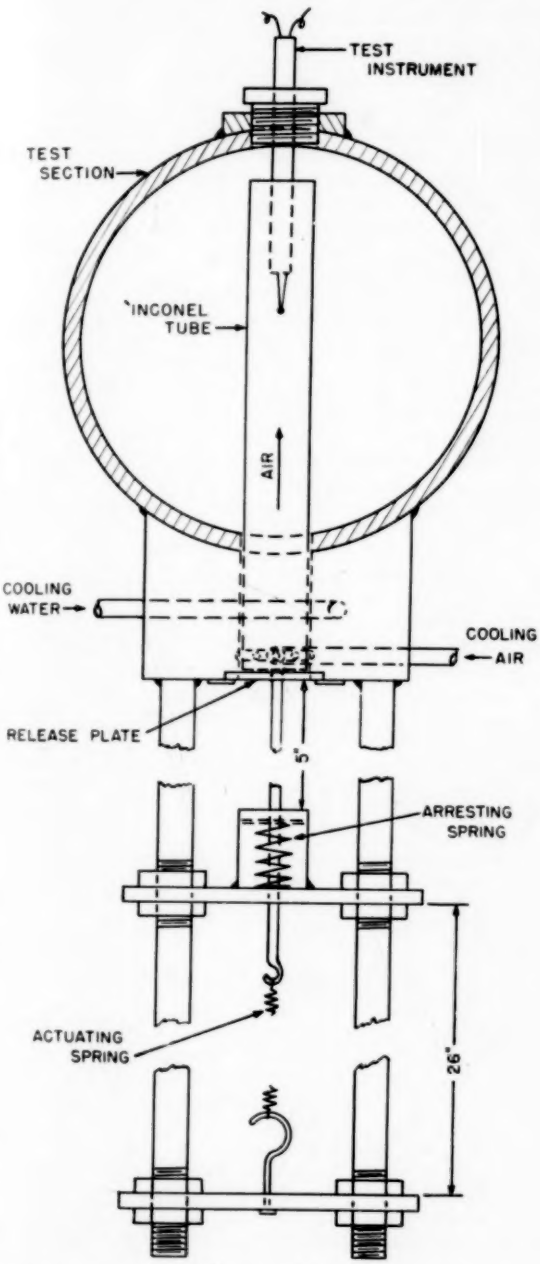

FIGURE 5. *Equipment for producing sudden changes in temperature.*

TABLE 1. *Characteristic times of bare thermocouple junctions*

Gage number of thermocouple wire	Characteristic time at mass flow rates (lb/sec ft^2) of—		
	2.2	4.2	6.8
	sec	*sec*	*sec*
18	2.2	1.6	1.2
22	1.3	1.0	0.8
28	0.54	0.41	.34
36	.17	.12	.10
40	.11	.08	.07
44	.07	.05	.04

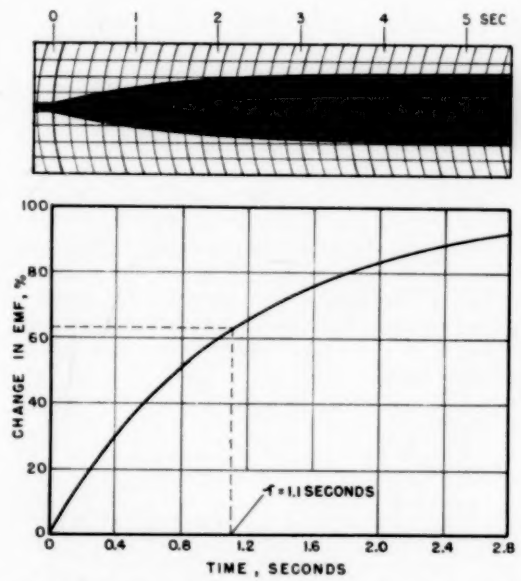

FIGURE 6. *Typical record of the response of a Chromel-Alumel thermocouple.*

worth while. Some typical results obtained for bare, untwisted Chromel-Alumel thermocouples are presented in table 1, merely to give an idea of the magnitudes of the characteristic times with which the designer of control equipment for jet engines must cope.

Examination of all the results obtained to date indicates that the effect of flow rate upon characteristic time is approximately the same for all instruments studied, and that the ratio of the characteristic times of any instrument at two flow rates is approximately proportional to the square root of the ratio of these mass flow rates.

From many measurements of characteristic time under experimental conditions, such that the principal heat transfer was by forced convection, it is apparent that the value observed for any particular instrument is essentially independent of the temperatures employed during the tests. Typical results illustrating this fact are shown in table 2. These results are for a No. 20 gage bare thermocouple, the junction of which was a bead of approximately twice the wire diameter, and for a mass flow rate of 6 lb/sec ft^2.

TABLE 2. *Effect of operating temperature range on observed characteristic time, when heat is transferred by forced convection*

Run	Temperature		Characteristic time (flow rate 6 lb/sec ft^2)
	T_1	T_2	
	°F	°F	sec
1	300	1,000	1.1
2	400	1,200	1.2$_5$
3	600	1,200	1.1
4	700	1,500	1.2
avg	-----	----- --	1.1$_6$

The present equipment was designed primarily to subject the test junction to a sudden increase in temperature. By heating the test junction electrically while it was surrounded by the Inconel tube, then subjecting it to a stream of unheated air, characteristic times for the cooling process were obtained. Results agree with the prediction of the theory that applies for the case of heat transfer by forced convection, in that the characteristic time at a given mass flow is the same, whether the instrument be heated or cooled.

VI. Conclusion

The characteristic time of a temperature-sensing element immersed in the gas stream of a turbine-type engine depends primarily on the mass flow rate, and is essentially independent of the temperature level, of the temperature difference over which it is measured, and of the direction of net heat flow. This is true primarily because the rate of heat transfer by forced convection is large compared with the rates of transfer by radiation and natural convection. Where radiation and natural convection are the controlling processes, the rate of response of an object to sudden changes in temperature depends greatly upon the actual temperatures involved. In the case of heat transfer by radiation only, different rates are obtained when the object is heated and cooled between the same two temperatures. Therefore, values of characteristic time that are significant in engine applications must be determined with equipment in which the mass flow rates approximate those prevailing in engines.

Apparatus of this kind is described in detail, and some typical results are presented. Results with a wide variety of instruments confirm the conclusions stated above.

The rate of response of a temperature-sensing element in the gas stream of an aircraft gas turbine varies greatly with the altitude of flight and with the engine speed. Adequate protection of the power plant requires a control system that will respond with the necessary rapidity during starting of the engine and at the flight ceiling. If this is accomplished, then the protection provided at all other altitudes and engine speeds should be more than adequate.

VII. References

[1] D. R. Harper, 3d, Thermometric lag, Bul. BS **8,** 659 (1912) S 185.

[2] Rudolf Beck, Thermometric time lag, Trans. Am. Soc. Mech. Engrs. **63,** 531 (1941).

[3] A. J. Hornfeck, Response characteristics of thermometer elements, Trans. Am. Soc. Mech. Engrs. **71,** 121 (1949).

[4] H. B. Hendrickson, Thermometric lag of aircraft thermometers, thermographs, and barographs, BS J. Research **5,** 695 (1930) RP222.

[5] W. N. Goodwin, Jr., Response time and lag of a thermometer element in a protecting case, Am. Inst. Elec. Engrs. **64,** 665 (1945).

[6] W. H. McAdam, Heat transmission, p. 241, 2d ed. (McGraw-Hill Book Co., Inc., New York, N. Y., 1942).

[7] Max Jacob, Heat transfer **1,** 529 (John Wiley & Sons, Inc., New York, N. Y., 1949).

WASHINGTON, June 7, 1950.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2137

Dielectric Constants of Aqueous Solutions of Dextrose and Sucrose

By Cyrus G. Malmberg and Arthur A. Maryott

The dielectric constants of aqueous solutions of dextrose and of sucrose prepared from NBS Standard Samples were determined at 20°, 25°, and 30° C for concentrations up to 50 and 60 percent by weight, respectively. A bridge method was used, and the sources of error involved are considered. The results, relative to a selected value for water, are considered to be accurate to 0.05 percent or better and are expressed in terms of an empirical equation for the purposes of interpolation.

I. Introduction

Examination of the literature reveals rather poor agreement in the data reported for the dielectric constants of aqueous solutions of dextrose [1 to 3][1] and of sucrose (static values) [4 to 18], (high-frequency data) [19 to 27]. Consequently, it seemed desirable to redetermine their static dielectric constants over a range of concentration in the vicinity of room temperature. Inasmuch as highly purified samples of dextrose and sucrose are available as standard samples at the National Bureau of Standards, these solutions may prove useful as secondary standards of dielectric constant having values intermediate between water and some of the more polar organic solvents.

For determinations of dielectric constants of media having an appreciable but not excessively large conductance, a bridge method employing relatively low and variable frequencies, as in this investigation, has certain advantages over resonance methods with regard to flexibility, accuracy, and the analysis and evaluation of errors. Errors such as arise from electrode polarization or lead inductance may be evaluated from measurements at various frequencies or often made negligible by the proper choice of frequency. Connection and calibration errors can be readily minimized by employing a cell of high capacitance. Perhaps the principal disadvantage is the difficulty sometimes experienced in maintaining sufficient stability of cell resistance to permit sensitive capacitance balance at lower frequencies.

II. Apparatus

The a-c bridge was of the equal ratio capacitance-conductance type, as indicated in figure 1. The ratio arms, R_1 and R_2, were 1,000-ohm woven resistors (Leeds & Northrup Type Std. 697–D) matched to within 0.001 percent in resistance. The measuring arm of resistance R_4 and capacitance C_4 consisted of a 100,000-ohm Leeds & Northrup resistance box (Cat. No. 4764–S) in parallel with a calibrated variable condenser, General Radio Co. Type 722–D. To extend the range of capacitance, General Radio Company's type 505 mica condensers were plugged to the terminals of the precision con-

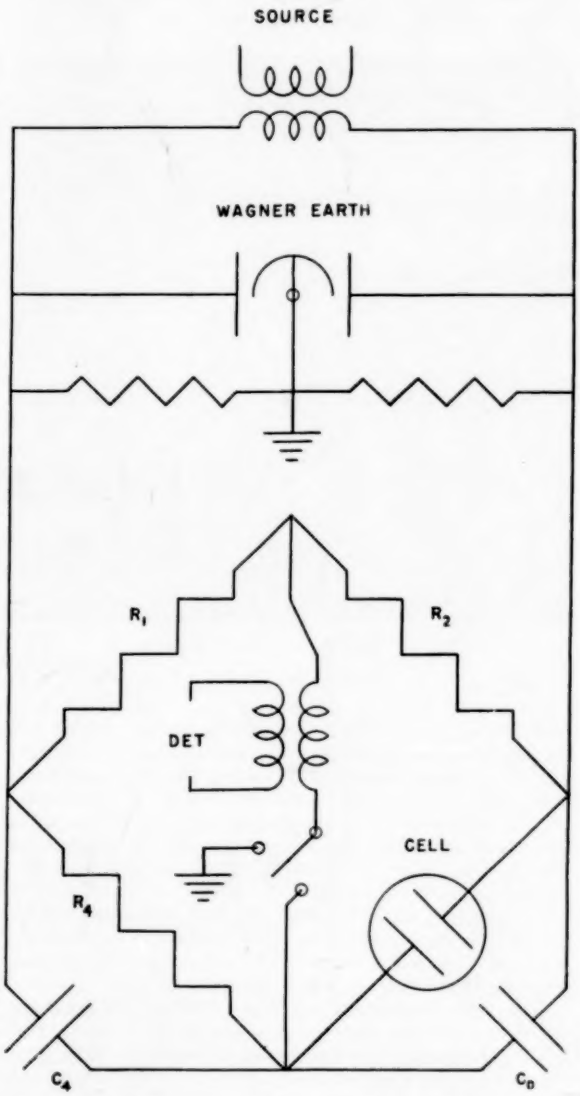

FIGURE 1. *Schematic circuit arrangement of capacitance-conductance bridge.*

[1] Figures in brackets indicate the literature references at the end of this paper.

denser as needed. The mica condensers were calibrated in terms of the precision condenser.

The unknown arm contained the cell and a relatively small variable capacitor, C_0, in parallel. C_0 remained fixed throughout the measurements and served only to provide a convenient "zero" setting for C_4. As no suitable three-terminal cell was available, two unshielded cells of convenient design were employed, each consisting of two concentric cylinders of bright platinum mounted in a glass container. Cell A had a replaceable capacitance of about 30 $\mu\mu$f and has been described previously [28]. Cell B, designed principally for measurements of conductivity, has a replaceable capacitance of approximately 60 $\mu\mu$f and differed from cell A in that it had parallel rather than concentric leads. In addition to the electrode compartment, each cell contained a reservoir to accommodate excess liquid and to permit mixing by forcing the contents back and forth with compressed gas. The construction of cell B is shown in figure 2 with its reservoir dismounted. Concentric mounting of its electrodes was accomplished by sealing to glass beads corresponding pairs of the six small platinum wires, which were welded to each end of the cylinders. The beads were then sealed to the outer walls of the cell. Two of the wires were sealed through the beads to provide contact to the mercury leads.

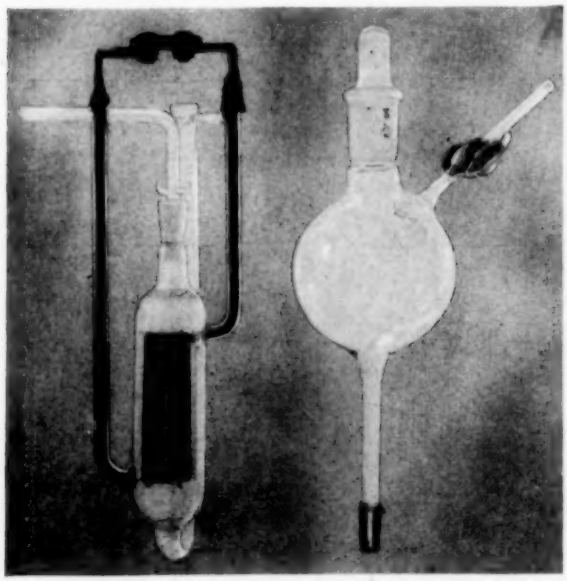

FIGURE 2. *Cell B with dismounted reservoir.*

Length of electrodes is 77 mm, separation 1 mm, and the diameter of the outer electrode 28 mm.

The cells were mounted in a reproducible position in the grounded thermostat and connected to the bridge terminals, using parallel leads of constant capacitance and inductance. Capacitance balance of the bridge was consistent to 0.1 $\mu\mu$f on replacement of the cell in the bath and reconnection to the bridge. The outer electrode was maintained at earth potential through the use of a Wagner earth. The use of this device along with appropriate shielding of the other branches of the bridge was effective in eliminating or fixing the influence of earth admittance and interbranch couplings associated with these branches and the corners of the bridge. Consequently, with the exception of a residual effect due to distributed interelectrode admittance to earth, the measured impedances of the cells were independent of the influence of earth admittances. On the basis of the construction of these cells and their measured earth capacities, this effect was estimated to be no greater than 0.2 $\mu\mu$f and therefore negligible.

A tuned circuit oscillator variable from 30 cycles to 200 kc was used as the source. A three-stage resistance-coupled amplifier and headphones were used as a detector up to 12 kc. At higher frequencies the amplifier was replaced by a Sargent model 11–XA regenerative receiver tuned to give any convenient audible beat-frequency. Throughout the measurements the sensitivity of capacitance balance was 0.1 $\mu\mu$f or better. Both source and detector were isolated from the bridge proper by shielded transformers.

Measurements were made at 20°, 25°, and 30°C with the cell in an oil thermostat controlled to within 0.01°.

III. Materials and Solutions

Solutions of sucrose of the desired concentrations were prepared on a weight basis from NBS Standard Sample 17 and conductivity water. Solutions of dextrose were prepared similarly from Standard Sample 41 and were then allowed to stand at least 4 hours in order to establish equilibrium with respect to mutarotation. Freshly prepared solutions of dextrose (α-form) had decidedly higher dielectric constants than the equilibrium mixtures. Conductivity water was obtained from ammonia-free distilled water by redistillation. It had a conductivity of about 10^{-7} mho before exposure to the atmosphere.

Initial attempts to determine the dielectric constants of the solutions of dextrose were unsuccessful because of the very rapid and continued increase in conductance that occurred when the solutions were introduced into the cell. Overnight the conductance sometimes increased by as much as a hundredfold. As this change was accompanied by a pronounced decrease in pH and occurred only when the solution was in contact with the electrodes, catalytic oxidation was suspected. Replacement of air by hydrogen eliminated this instability. Consequently, hydrogen was used to flush out the cell, to saturate each solution, and to transfer the solution from the flask to the cell. The same procedure was adopted for the solutions of sucrose because of the advantage of lower conductances gained by removal of dissolved CO_2. The specific conductivities of the solutions were 3×10^{-6} mho or lower in all cases.

IV. Procedure

The dielectric constants were determined relative to water at 25°, using the value reported by Wyman

[29], namely, 78.54. The replaceable vacuum capacitance of the cell, C_V, was obtained from the difference in capacitance of the cell first in air and then when filled with water, $C_W - C_A$, respectively, according to

$$C_V = (C_W - C_A)/77.54 \quad (1)$$

The dielectric constant of the unknown medium, ϵ_X, is then

$$\epsilon_X = 78.54 - \Delta C_X/C_V, \quad (2)$$

where ΔC_X is the change in capacitance of the cell on replacing the unknown medium by water at 25°.

Before the experimental procedure is discussed in any detail, the more significant sources of error will be considered briefly. These errors are associated for the most part with residual impedances in the arms of the bridge network. To a sufficient degree of approximation, the bridge may be simulated by the network illustrated in figure 3. Reactance in the ratio arms is represented by the equivalent parallel capacitances, C_1 and C_2. L_M and L_X are, respectively, the self-inductances of the leads to the precision condenser and to the cell. The effective series inductance of the resistance box is represented by L_4 and the resistance of the cell by R_X. During the course of the measurements, C_M and R_4, and consequently L_4, are varied. The observed change in capacitance of the precision condenser, ΔC_M, is not generally equal to ΔC_X but related according to

$$\Delta C_M = \Delta C_x + \Delta C_a + \Delta C_b + \Delta C_c + \Delta C_d \quad (3)$$

where, on neglecting terms of higher order,

$$\Delta C_a = (C_1 - C_2)(R_1/R_4 - R_1/R_4),$$

$$\Delta C_b = \frac{L_4}{R_4^2} - \frac{L_x}{R_4^2} - \frac{L_4'}{(R_4')^2} + \frac{L_x}{(R_4')^2},$$

$$\Delta C_c = \omega^2(L_x - L_M)(C_M + C_M)\Delta C_M,$$

ω being the angular frequency. The primed symbols refer to solution in the cell and the unprimed to water. ΔC_a represents the effect of reactive asymmetry of the ratio arms; ΔC_b, the effect of inductance in series with resistance in the measuring and unknown arms; ΔC_c, inductance in series with capacitance in the measuring and unknown arms, and ΔC_d the change in capacitance attributable to electrode polarization.

The effect of polarization capacitance, C_d, was found to be proportional to ω^{-n}, where $n<2$ and constant for a given cell and solution. On the other hand, C_c is proportional to ω^2, so the influence of these frequency dependent errors was readily determined by varying the frequency between 3 and 96 kc. Actually, in this investigation it was not necessary to extrapolate out these errors, as they became quite negligible at some intermediate frequency where C_M reached its minimum value. This frequency was usually 48 kc for the solutions of dextrose (cell A) and 24 kc for the solutions of sucrose (cell B).

The frequency-independent errors, expressed as the sum $\Delta C_a + \Delta C_b$, were eliminated by a procedure that was essentially equivalent to making all measurements at the same resistance. With the cell filled with water at constant temperature, its resistance was varied to correspond to the resistances of the solutions by increasing or decreasing the CO_2 content. In this manner the change in measured capacitance accompanying a given change in cell resistance was determined. The resistances of the solutions were usually between 1,000 and 3,000 ohms, so that correspondence of the 1,000- and 100-ohm dials only was needed. These corrections were relatively small and amounted to 0.1 unit or less in dielectric constant. The maximum concentration of carbon dioxide introduced into the water during the course of this determination is estimated to be about 2×10^{-4} mole/liter. Considering both the ionized and unionized portions, estimates of the effect of this amount of carbon dioxide on the dielectric constant indicate a negligible change, of the order of 0.003 unit or less.

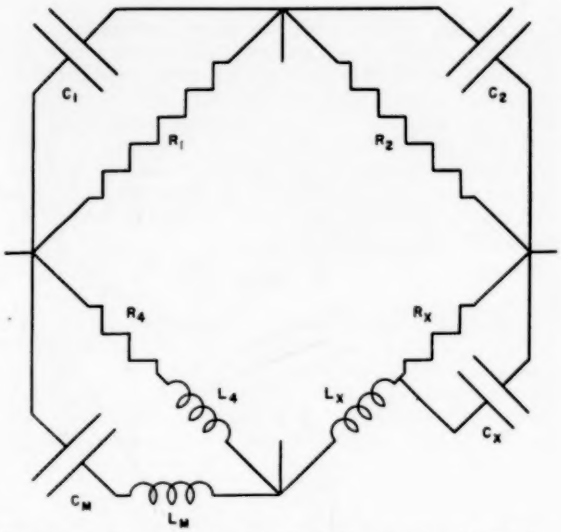

FIGURE 3. *Simulated circuit of the bridge showing distribution of effective impedances in the arms of the bridge.*

Values of C_V for both cells were determined with a known fixed resistor, R_A, directly across the arm containing the cell both when empty and when filled with water at 25°C. Because of this shunt (since R_x is no longer equal to R_4), the relation of ΔC_x to ΔC_M given in eq 3 is altered with regard to the term ΔC_b with the result that now

$$\Delta C_b = \frac{L_4}{R_4^2} - \frac{L_x}{R_x^2} - \frac{L_4'}{R_A^2},$$

where $R_x = R_4R_A/(R_A - R_4)$. Since for corresponding values of R_4, the difference between this value of ΔC_b and that for direct measurement involves only terms in L_x and known values of R, L_X was evaluated, and the experimentally determined values of $\Delta C_a + \Delta C_b$ were corrected as needed for use in calculating ΔC_X to obtain C_V.

Although neither cell contained insulating supports between the electrodes, their construction did not preclude some nonlinearity in the variation of capacitance with dielectric constant of the medium. That no detectable error was encountered in this range of dielectric constant was indicated by intercomparisons between the two cells and by a determination of the variation of C_V for cell B by using liquids, the dielectric constants of which were determined by an absolute method.

V. Results

The experimental data, including those for water at 20° and 30° C, are listed in table 1. Inasmuch as the actual concentrations were very close to even values of weight percent, the dielectric constants have been interpolated to even values. All of the data for the dextrose solutions can be expressed by the equation,

$$\epsilon\ (\text{dextrose}) = \epsilon_w - [0.227 - 1.2(10^{-3})(t-25)]C - [1.32(10^{-3}) + 1.2(10^{-5})(t-25)]\ C^2 - [3.9(10^{-6}) - 2.8(10^{-7})(t-25)]\ C^3,$$

where ϵ(dextrose) and ϵ_w are, respectively, the dielectric constants of the dextrose solution and of water at the temperature, t, in degrees Celsius, and C is the weight percent of dextrose. This equation fits the data of table 1 with an average deviation of 0.02 and a maximum deviation of less than 0.04 unit in dielectric constant. For the solutions of sucrose, the corresponding equation,

$$\epsilon\ (\text{sucrose}) = \epsilon_w - 0.226C - [6.75(10^{-4}) - 1.5(10^{-5})(t-25)]C^2 - [1.09\ (10^{-5}) + 4(10^{-8})(t-25)]\ C^3,$$

fits the data with an average deviation of 0.01 and a maximum deviation of less than 0.03 unit. Exclusive of any uncertainty in the value selected for water, these data are probably accurate to 0.05 percent or better.

TABLE 1. *Dielectric constants of aqueous dextrose and sucrose solutions*

Dextrose	Dielectric constant at—		
	20° C	25° C	30° C
wt percent			
0	80.38	78.54	76.76
5	79.17	77.37	75.64
10		76.14	
15	76.56	74.80	73.11
20		73.43	
30	72.13	70.46	68.82
40	68.73	67.11	65.56
50	64.90	63.39	61.91
Sucrose			
wt percent			
0	80.38	78.54	76.76
10	78.04	76.19	74.43
20	75.45	73.65	71.90
30	72.64	70.86	69.13
40	69.45	67.72	66.05
50	65.88	64.20	62.57
60	61.80	60.19	58.64

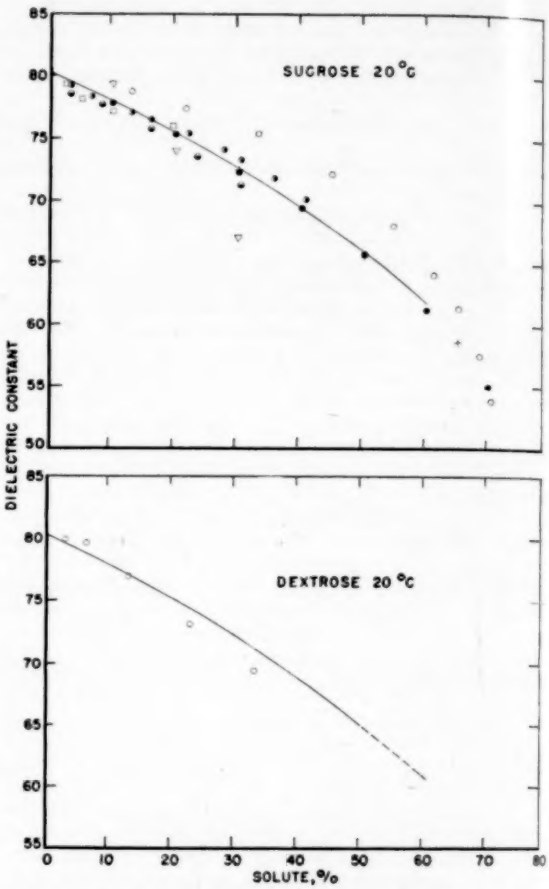

FIGURE 4. *Comparison of dielectric constant data reported on aqueous solutions of dextrose and sucrose.*

●: Kockel—10, 20, 30, and 40 percent; 600 kc. Davies—10, and 20 percent; 1.5 to 4 mc. Kniekamp—10 and 20 percent; 566 kc. Åkerlöf—50, 60, and 70 percent; 2 mc. These data are plotted under a common symbol, as the differences in value of ϵ reported could not be resolved for the purposes of this graph. ○, Berliner and Rüter—3.3 mc; ◑, Harrington—1 mc; □, Kistler—2.1 mc; ▽, Furth—400 mc; ◒, Astin—2.7 mc; +, Slevogt—3.1 mc; ——, present data (points fall uniformly on the smooth curve within the precision of the scale of plotting).

A comparison of these results at 20° C with those of some other workers after correction to the same dielectric constant for water is given in figure 4. The curves represent the present data and that for dextrose is extrapolated to 60 percent for comparative purposes. Of the data on dextrose, only those of Berliner and Rüter obtained at 3.3 mc are comparable to the present data as representing static values, and the agreement is poor. For the data on sucrose the best agreement is shown by those of Kockel, Åkerlöf, Kniekamp, Slevogt (at 3.1 mc), and to a lesser extent by those of Davies as reported by Lattey, Gatty, and Davies. These values are, in general, lower than the present data by from 0.1 to 1 percent and were obtained by resonance methods.

Berliner and Rüter also used a resonance method. Astin developed and used a method employing voltage resonance that was subsequently adopted by Åkerlöf. The values reported by Harrington were obtained by means of a bridge method at 1 mc, and it is uncertain what method Kistler used at 2.1 mc. The measurements of Furth (400 mc) and others [19 to 27] were made at very high frequencies, so the low values reported are undoubtedly attributable to anomalous dispersion and are not comparable to static values.

Consideration of the data available on static values suggests that the major descrepancies in the values reported are in large part due to high conductivity as, in general, it appears that precautions have not been taken to obtain and maintain low conductivities for these solutions. To a lesser extent it is also probable that frequency-dependent errors are associated with many of these values. The extent to which these factors may modify the suitability of a method of measurement is exemplified by Kniekamp's [9] study and Hartshorn's analysis [30] of resonance methods.

VI. References

[1] E. Berliner and R. Rüter, Kolloid-Z. **47,** 251 (1929).
[2] R. Furth, Ann. Physik. **70,** 63 (1923).
[3] G. Kreinina, J. Exptl. Theoret. Phys. (USSR) **15,** 208 (1945).
[4] G. Åkerlöf, J. Am. Chem. Soc. **54,** 4125 (1932).
[5] A. V. Astin, Phys. Rev. **34,** 300 (1929).
[6] E. Berliner and R. Rüter, Kolloid-Z. **47,** 251 (1929).
[7] E. H. Harrington, Phys. Rev. **8,** 581 (1916).
[8] S. S. Kistler, J. Phys. Chem. **35,** 815 (1931).
[9] H. Kniekamp, Z. Physik **51,** 95 (1928).
[10] L. Kockel, Ann. Physik **77,** 417 (1925).
[11] E. Landt, Centr. Zuckerind. **44,** 723 (1936).
[12] R. T. Lattey, Phil. Mag. **41,** 829 (1921).
[13] R. T. Lattey, O. Gatty, and W. G. Davies, Phil. Mag. **12,** 1019 (1931).
[14] H. Possner, Ann. Physik **6,** 815 (1930).
[15] G. Scatchard, J. Am. Chem. Soc. **48,** 2026 (1926).
[16] C. Schreck, Physik. Z. **37,** 156 (1936).
[17] K. E. Slevogt, Ann. Physik. **36,** 141 (1939).
[18] P. Walden and O. Werner, Z. physik. Chem. **129,** 405 (1928).
[19] M. Von Ardenene, O. Groos, and G. Otterbein, Physik. Z. **37,** 533 (1936).
[20] W. Dahms, Ann. Physik **26,** 177 (1936).
[21] O. Dobenecker, Ann. Physik **17,** 699 (1933).
[22] P. Drude, Z. physik. Chem. **23,** 267 (1897).
[23] A. Duebner, Ann. Physik **84,** 429 (1927).
[24] R. Furth, Ann. Physik **70,** 63 (1923).
[25] R. Keller, Kolloid-Z. **29,** 193 (1921).
[26] G. Kreinina, J. Exptl. Theoret. Phys. (USSR) **15,** 208 (1945).
[27] P. Wenk, Ann. Physik **17,** 679 (1933).
[28] A. A. Maryott, J. Research NBS **38,** 527 (1947) RP1794.
[29] J. Wyman, Phys. Rev. **35,** 623 (1930).
[30] L. Hartshorn, Radio frequency measurements by bridge and resonance methods (John Wiley & Sons, Inc., New York, N. Y., 1940).

WASHINGTON, April 13, 1950

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2138

Response of Accelerometers to Transient Accelerations

By Samuel Levy and Wilhelmina D. Kroll

Curves and tables are shown for the response of accelerometers to transient exciting accelerations. Three types of acceleration-time relations are considered. When plotted, they have square, triangular, and half-sine-wave shapes. The natural periods of the accelerometers for which the computations were made were approximately one, one-third, and one-fifth of the duration of the acceleration pulse. The damping coefficients of the accelerometers were 0, 0.4, 0.7, and 1.0 times the critical values. It is indicated that, to obtain an accuracy of better than 5 percent of the peak acceleration in measuring acceleration pulses having the general characteristics of the triangular or sinusoidal pulses, an accelerometer must have a natural period of about one-third the duration of the acceleration pulse, and a damping constant of about 0.4 to 0.7 of the critical value.

I. Introduction

Accelerometers are widely used to measure oscillatory and transient vibrations.

The fidelity with which these instruments respond in the case of oscillatory stimuli has been thoroughly studied [1, pp. 61 to 70]. It is found that, when the damping is between 0.6 and 0.7 of the critical value and the natural period of the accelerometer is less than about half of the period of the applied acceleration, the accuracy is satisfactory.

In the case of excitation of the accelerometer by a transient vibration, only scattered information is available regarding the reliability of the response obtained. Weiss [2] gives the response to a triangular pulse of acceleration for an accelerometer whose natural period is 0.3 the duration of the pulse and whose damping is 0, 0.3, and 0.7 of the critical value. He also gives the response to a suddenly applied constant acceleration for accelerometers with a damping ratio of 0, 0.3, 0.7, and 1.0 times the critical value. Welch [3] has determined, on the Westinghouse transient analyzer, the response to several kinds of impulses of a 50 c/s single-degree-of-freedom shock measuring instrument having various amounts of damping. On the basis of these scattered data, and information for undamped accelerometers derived by Frankland [4], Biot and Bisplinghoff [5], and others, it has been common practice to assume that an accelerometer will be acceptable in a given

application if its damping is 0.6 to 0.7 of the critical value and if its natural period is less than about half of the duration of the acceleration pulse.

The curves presented in this report were computed to obtain more systematic information regarding the accuracy of damped accelerometers in measuring transient phenomena.

II. Theory

The usual accelerometer is a single-degree-of-freedom mechanical system. Such a system is shown in figure 1. Means are provided to indicate the relative motion x of the internal mass with respect to the frame. This relative motion is taken as a measure of the acceleration, d^2y/dt^2, of the frame.

The equation of motion for the mass m, figure 1, is

$$m\frac{d^2z}{dt^2}+c\left(\frac{dz}{dt}-\frac{dy}{dt}\right)+k(z-y)=0. \tag{1}$$

With

$$z=x+y, \tag{2}$$

eq 1 becomes

$$m\frac{d^2x}{dt^2}+c\frac{dx}{dt}+kx=-m\frac{d^2y}{dt^2}, \tag{3}$$

$$\frac{d^2x}{dt^2}+\frac{c}{m}\frac{dx}{dt}+\frac{k}{m}x=-\frac{d^2y}{dt^2}. \tag{4}$$

We wish to know how faithfully the response x of the accelerometer reproduces the time history of the applied acceleration d^2y/dt^2 for pulses of acceleration of finite duration and arbitrary shape. To give the analysis a wider range of usefulness, eq 4 is written in dimensionless form by making the following substitutions:

$$\left.\begin{aligned} a&=\left(\frac{d^2y}{dt^2}\right)\Big/\left(\frac{d^2y}{dt^2}\right)_{max}\\ \tau&=t/T\\ \xi&=-kx/m\left(\frac{d^2y}{dt^2}\right)_{max}\\ D&=c/2\sqrt{mk}=c/c_c\\ R&=2\pi\sqrt{m/k}/T \end{aligned}\right\} \tag{5}$$

where

$c_c=2\sqrt{mk}$, critical value of damping coefficient,

T=duration of acceleration pulse to be measured,

$2\pi\sqrt{m/k}$="undamped" period of accelerometer,

$\left(\frac{d^2y}{dt^2}\right)_{max}$=peak value of acceleration.

Substituting eq 5 into eq 4 gives

$$\left(\frac{R}{2\pi}\right)^2\frac{d^2\xi}{d\tau^2}+\frac{RD}{\pi}\frac{d\xi}{d\tau}+\xi=a. \tag{6}$$

For a relatively high frequency accelerometer, R is a small number. Under these circumstances, the first two terms in eq 6 become negligible, and the dimensionless response ξ is equal to the dimensionless acceleration a. As R becomes larger, the first and second terms start to have an effect. The primary effect of the second term is to introduce a time lag between the response ξ and the acceleration a. The primary effect of the first term is to tend to make the response ξ oscillate in value above and below the value of the acceleration a.

III. Results

Equation 6, giving the relation between the dimensionless responses ξ and the dimensionless acceleration a, was integrated numerically for three values of the natural period ratio having approximately the values, R=1, 1/3, 1/5; for four values of the damping ratio, D=0, 0.4, 0.7, 1.0; and for the three time-histories of acceleration pulse shown in figure 2. Numerical integration, instead of direct integration, was used to give results that could be plotted directly. Small variations from the nominal values of R were used for convenience in computing. These values of R are given in table 1. A spot check of the results was made using the analytical solution of eq 6.

The numerical integration was carried out using a time increment of $1/(20\pi)$ times the natural period of the accelerometer. Eight decimal figures were used in the computation.

The results are plotted in figures 3 to 11. Figures 3, 4, and 5 give the response to a sinusoidal pulse of acceleration. Figure 3 gives the response when the natural period is about equal to the duration of the acceleration pulse. Figures 4 and 5 give similar results with the natural period about one-third and one-fifth, respectively, of the duration of the acceleration pulse. In each figure, the dimensionless applied acceleration, a, is shown by a dotted line; the response, ξ, with the damping ratio D=0 by curve 1; with D=0.4 by curve 2; with D=0.7 by curve 3; and with D=1.0 by curve 4.

Figures 6, 7, and 8 show the response to a triangular pulse of acceleration, and figures 9, 10, and 11 show the response to a rectangular pulse. In each figure, the set of curves brings out the effect of varying only the damping ratio D.

IV. Discussion

It is evident from an inspection of the figures that for none of the accelerometers considered does the time history of the dimensionless response ξ coincide with the time history of the dimensionless accelera-

TABLE 1.—*Error of accelerometers and acceleration pulses considered*

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Error				Plotted	
R	D	$\xi_{max}-a_{max}$	$\lvert\xi-a\rvert_{max}$	After shift $\lvert\xi-a\rvert_{max}$	$\Delta\tau$	Figure	Curve
HALF-SINE-WAVE PULSE							
1.014	0	0.74	1.36	1.36	0.17	3	1
1.014	0.4	.15	0.62	0.24	.22	3	2
1.014	.7	−.05	.67	.18	.25	3	3
1.014	1.0	−.18	.65	.24	.29	3	4
0.338	0	.17	.35	.35	.07	4	1
.338	0.4	.02	.20	.06	.05	4	2
.338	.7	.00	.24	.06	.07	4	3
.338	1.0	−.03	.32	.10	.11	4	4
.203	0	.10	.20	.20	.01	5	1
.203	0.4	.07	.12	.05	.02	5	2
.203	.7	.00	.15	.03	.04	5	3
.203	1.0	−.01	.20	.06	.06	5	4
TRIANGULAR PULSE							
1.014	0	0.51	1.29	1.29	0.15	6	1
1.014	0.4	.00	0.56	0.21	.21	6	2
1.014	.7	−.20	.51	.20	.22	6	3
1.014	1.0	−.32	.57	.32	.25	6	4
0.338	0	.17	.43	.43	.04	7	1
.338	0.4	.00	.18	.06	.07	7	2
.338	.7	−.06	.17	.06	.08	7	3
.338	1.0	−.11	.22	.12	.08	7	4
.203	0	.10	.25	.25	.01	8	1
.203	0.4	.00	.11	.03	.04	8	2
.203	.7	−.04	.10	.04	.05	8	3
.203	1.0	−.07	.13	.07	.07	8	4
SQUARE PULSE							
1.014	0	1.00	1.00	1.00	0.04	9	1
1.014	0.4	0.25	1.00	0.49	.19	9	2
1.014	.7	.05	1.00	.53	.22	9	3
1.014	1.0	−.01	1.00	.50	.28	9	4
0.334	0	1.00	1.00	1.00	.01	10	1
.334	0.4	0.25	1.00	0.50	.06	10	2
.334	.7	.05	1.00	.50	.08	10	3
.334	1.0	.00	1.00	.50	.10	10	4
.203	0	1.00	1.00	1.00	.02	11	1
.203	0.4	0.25	1.00	0.50	.03	11	2
.203	.7	.05	1.00	.50	.04	11	3
.203	1.0	.00	1.00	.50	.06	11	4

tion a. In many cases, however, the coincidence can be markedly improved by considering the response curves to be shifted a small distance to the left. They can also be improved, in those cases where oscillatory response is present, by fairing a line through the oscillatory response. Both of these methods of record improvement are commonly employed.

The errors of the various accelerometers for the acceleration pulses considered are given in table 1. In columns 1 and 2, respectively, are given the accelerometer characteristics: R, ratio of natural period to pulse duration; and D, ratio of damping constant to critical value.

In column 3, table 1, is given the difference between the maximum value of dimensionless response, ξ, and the maximum value of the dimensionless applied acceleration, a. The error varies from 0 to 100 percent.

In column 4, table 1, is given the largest absolute value of the difference $\xi-a$ where ξ and a are evaluated at the same dimensionless time. The error varies from a minimum of 10 percent to a maximum of 136 percent.

In column 5, table 1, is given the largest absolute value of the difference $\xi-a$ after shifting the ξ curve to the left by the amount $\Delta\tau$ given in column 6. The error in this case is typical of the usual way of interpreting accelerometer records. This error varies from a minimum of 3 percent to a maximum of 136 percent. If only accelerometers with damping are considered ($D>0$), the largest error is 32 percent when the accelerometer is subjected to accelerat on pulses of triangular or sinusoidal time his*tor*ies.

On the basis of the few cases investigated, an optimum value of damping is indicated to be between 0.4 and 0.7 of the critical value. It is also indicated that, to obtain an accuracy of better than 5 percent of the peak acceleration in measuring acceleration pulses having the general characteristics of the triangular or sinusoidal pulses, an accelerometer must have a natural period of less than about one-third the duration of the acceleration pulse.

Acknowledgment is due to the Bureau of Aeronautics, Navy Department, whose research projects on vibration pickups have provided the impetus for the work presented in this paper. The authors also extend thanks to L. W. Roberson and I. Smith for assistance in computing the many response curves and preparing the figures and table.

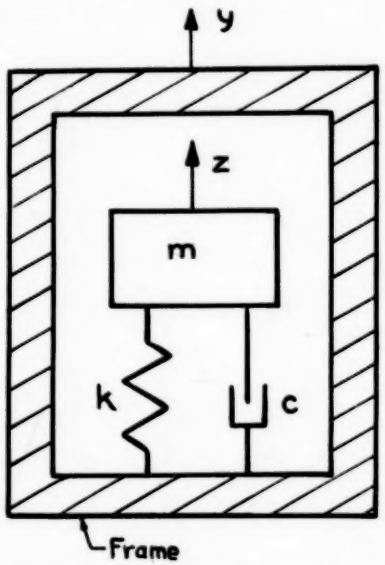

FIGURE 1. *Single-degree-of-freedom system representing accelerometer.*

Displacement of frame is y, displacement of internal mass is z, relative displacement of internal mass with respect to frame is $x=z-y$.

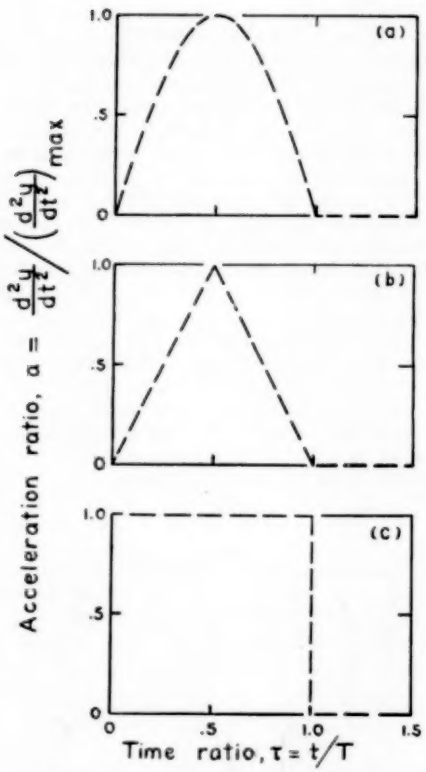

FIGURE 2. *Pulses of acceleration for which integration was carried out: (a) half-sine-wave pulse, (b) triangular pulse, (c) square pulse.*

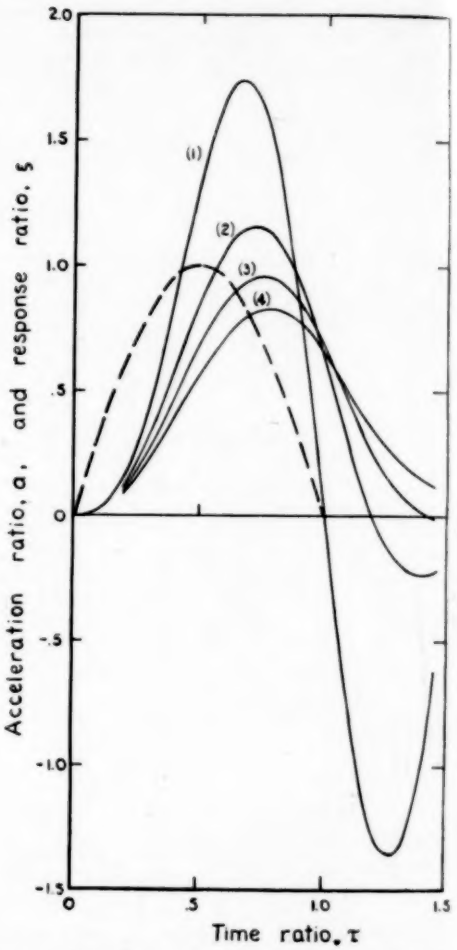

FIGURE 3. *Response to a half-sine-wave pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to the duration of the pulse, $R=1.014$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

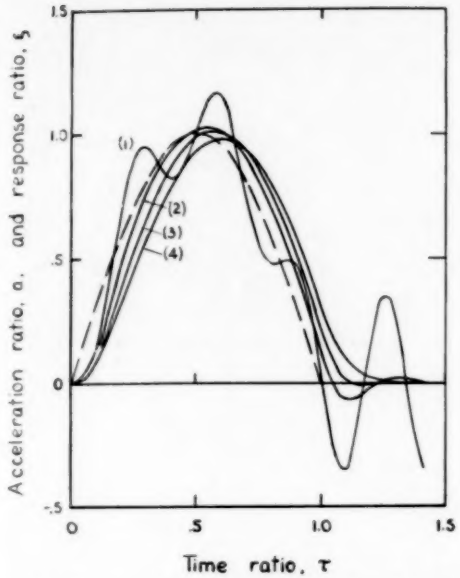

FIGURE 4. *Response to a half-sine-wave pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to one-third of the duration of the pulse, $R=0.338$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

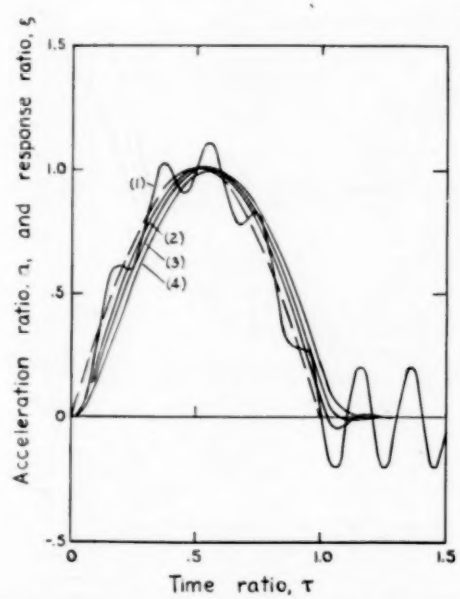

FIGURE 5. *Response to a half-sine-wave pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to one-fifth of the duration of the pulse, $R=0.203$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

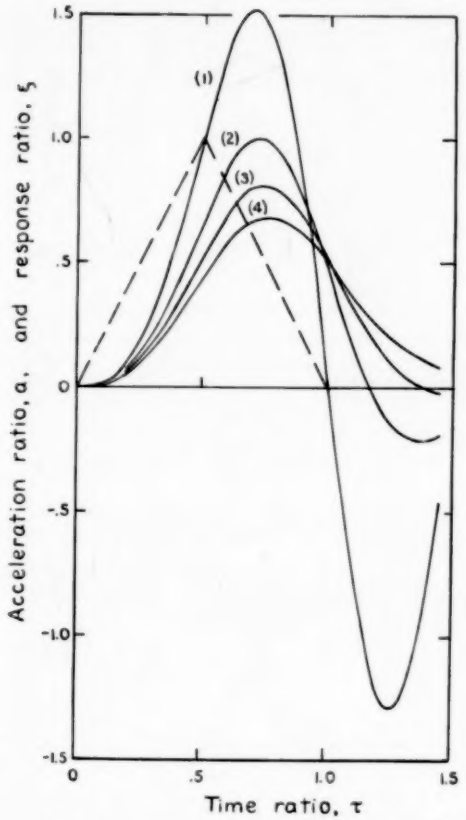

FIGURE 6. *Response to a triangular pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to the duration of the pulse, $R=1.014$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

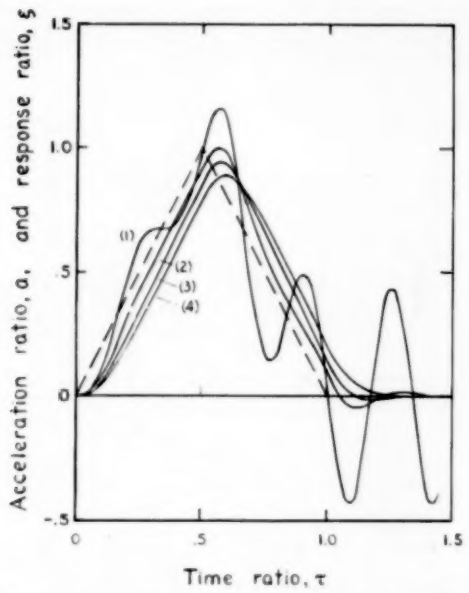

FIGURE 7. *Response to a triangular pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to one-third of the duration of the pulse, $R=0.338$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

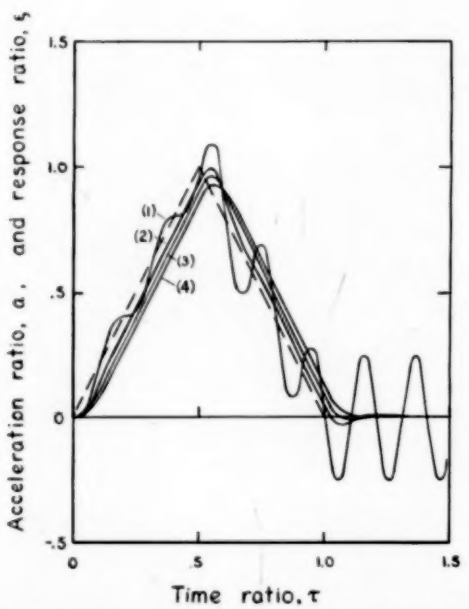

FIGURE 8. *Response to a triangular pulse of acceleration, dash curve, of an accelerometer whose natural period is about equal to one-fifth of the duration of the pulse, $R=0.203$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

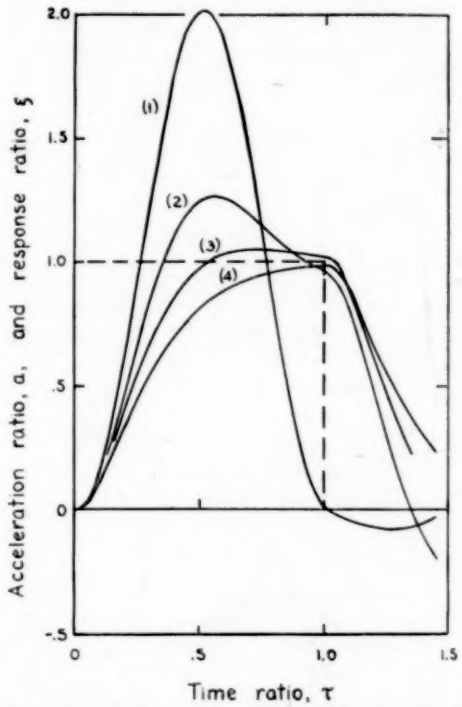

FIGURE 9. *Response to a square pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to the duration of the pulse, $R=1.014$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

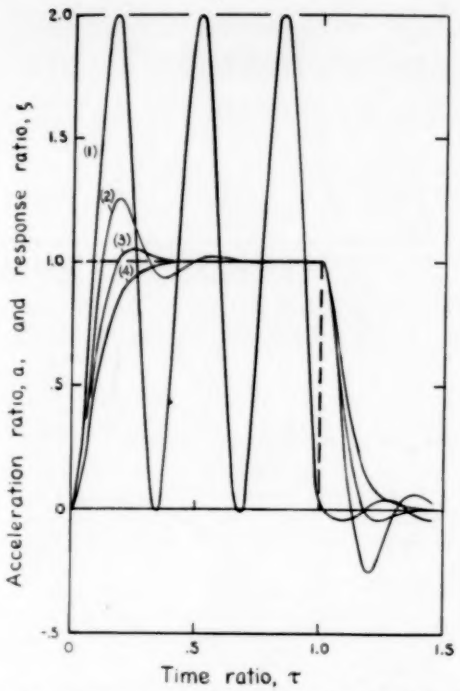

FIGURE 10. *Response to a square pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to one-third of the duration of the pulse, $R=0.334$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

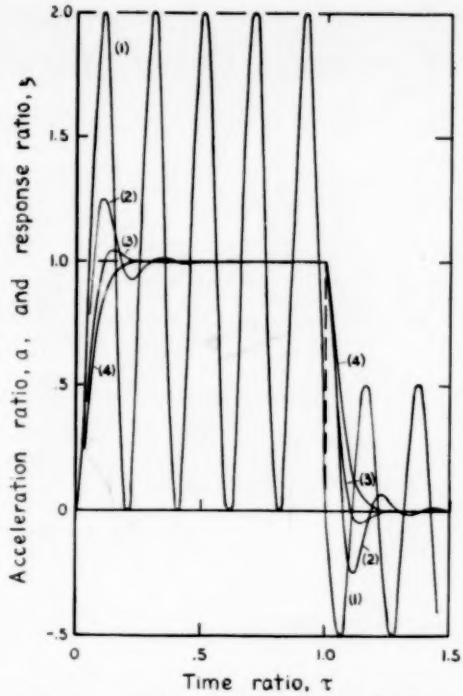

FIGURE 11. *Response to a square pulse of acceleration, dashed curve, of an accelerometer whose natural period is about equal to one-fifth of the duration of the pulse, $R=0.203$.*

Curve (1), damping coefficient zero, $D=0$; curve (2), damping coefficient 0.4 of the critical, $D=0.4$; curve (3), damping coefficient 0.7 of the critical, $D=0.7$; curve (4), damping coefficient equal to the critical, $D=1.0$.

V. References

[1] J. P. Den Hartog, Mechanical vibrations (McGraw-Hill Book Co., Inc., New York, N. Y., 1940).

[2] D. E. Weiss, Design and application of accelerometers, Proc. SESA, vol. IV, no. II, p. 89 (1947).

[3] W. P. Welch, A proposed new shock measuring instrument, Proc. SESA, vol. V, no. I, p. 39 (1947).

[4] J. M. Frankland, Effects of impact on simple elastic structures, Taylor Model Basin Rep. 481 (April 1942).

[5] M. A. Biot and R. L. Bisplinghoff, Dynamic loads on airplane structures during landing, NACA ARR No. 4H10 (Oct. 1944).

WASHINGTON, May 11, 1950.

Electrode Deterioration in Transmit-Receive Tubes[1]

By Judson C. French

An investigation of the failure of a modified 1B24 TR (transmit-receive) tube revealed the cause to be the short circuiting of the pulsed keep-alive cathode to its anode, due to the formation of an unusual deposit on the cathode insulation and on the nearby anode surfaces. Analysis of the deposit showed that it originated at the cathode. To facilitate the investigation, diodes were constructed in which such parameters as distributed capacity, cathode material, current density and wave-form, and water vapor content of the gas fill could be readily varied while maintaining constructional features of the modified 1B24 relevant to the problem. The study disclosed that the glow discharge normally present breaks into a momentary arc, which by local heating of the cathode frees the metal that forms the deposit. Critical factors in the processes were determined, and by their proper control the deposit could be reduced or under some circumstances eliminated.

I. Introduction

The transmit-receive, or TR, tube is an essential element of many types of modern microwave systems. In a radar-type circuit it permits a single antenna to be used alternately by the transmitter and by the receiver. Its dependable operational life is of great importance not only to the operation of the system as a whole, but to the life of the receiver that it protects from the high power pulses of the radar's own transmitter and those of neighboring systems. Incidence of such a pulse upon the tube causes a radio frequency discharge in its argon- and water-vapor-filled resonant cavity. The discharge detunes the cavity; the resulting mismatch reflects the incident signal and the sensitive receiver is thus protected from burnout. The protection is sufficient only if the discharge occurs rapidly enough to prevent excessive power leakage to the receiver. To ensure this rapid breakdown, a small supply of ions is maintained in a critical region of the cavity by a direct current (or pulsed) glow discharge, the so-called keep-alive discharge. Failure of the keep-alive discharge probably results in permanent damage to the receiver.

In a specialized version of a 1B24 TR tube, a deposit short circuiting the keep-alive discharge was found to form in a short time, ending the useful life of the tube prematurely. The investigation of the short circuit and remedial methods form the subjects of this paper. During the course of the investigation, it was found that breakdown of the glow discharge into an arc locally heated the cathode producing a metallic deposit that caused the short circuit. The extent of the deposit could be controlled by reduction of distributed capacity in the discharge circuit or by proper selection of cathode material. Current wave-form and density and the gas fill were also found to influence the deposit formation.

II. The TR Tube

The tube is illustrated in figure 1, a. The keep-alive cathode surface is the circular end of the tapered portion of a cylindrical kovar rod (fig. 1, b and c). The rod is insulated along its length by 7052 glass,

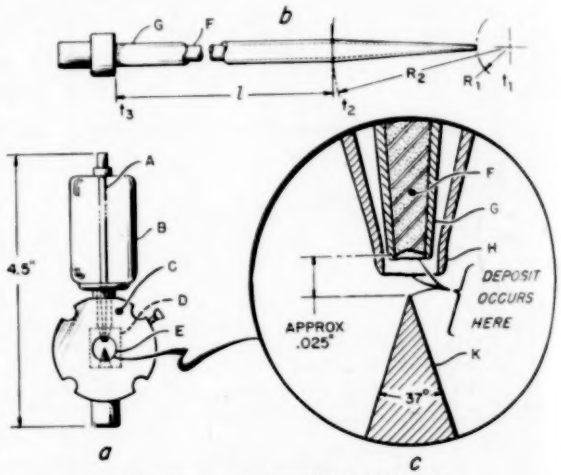

FIGURE 1. *The modified 1B24 tube.*

a, Exterior view of the tube; *A*, keep-alive cathode lead; *B*, gas reservoir; *C*, copper tube-body; *D*, internal resonant cavity containing cones between which radio frequency discharge occurs; *E*, window, coupling radio frequency energy to resonant cavity.

b, The keep-alive cathode assembly; *F*, the kovar cathode lead; *G*, glass insulation sealed to cathode lead.

c, Cross section of cone tips in cavity; *F*, the kovar cathode lead; *G*, glass insulation sealed to cathode; *H*, hollow copper cone concentric to cathode serving as one keep-alive anode; *K*, solid copper cone serving as a second anode. The radio-frequency discharge occurs between cones *H* and *K* to detune the cavity.

which reaches its minimum thickness of 5 mils around the edge of the cathode face. Separation of the outer edge of the glass insulation from the concentric hollow truncated copper cone anode is not uniform, but in any event is of the order of several mils (fig. 1, c).

During certain phases of operation of the tube, the copper anodes are grounded while a negative potential applied to the cathode from a high impedance source produces a square wave of current. Current density at the cathode is about 5 μa per square mil. This abnormal glow discharge takes place in an atmosphere of argon and water vapor at a total pressure of the order of 10 mm Hg. Electrical and visual indications were that the short circuit limiting the tube life appeared between the kovar cathode and the concentric cone anode.

[1] This work was sponsored by the Bureau of Ordnance of the Navy Department.

III. Preliminary Investigation of the Short Circuit

Short-circuited tubes were disassembled to permit visual study of the electrodes. This disclosed heavy deposits, as much as several mils thick, on the portions of the cones nearest the cathode and on the glass immediately surrounding the cathode surface. These deposits produced the short circuit at the point where they covered the edge of the insulating glass and bridged the gap between glass and concentric cone.

That the source of the material was the cathode itself was suggested by the fact that the cathode had been destroyed to a depth of 5 to 10 mils, leaving a bright surface rather uniformly covered with small pits. The pit diameters were of the order of 10^{-4} in. The suggestion was confirmed by spectroscopic analysis of samples of the deposited crust, which showed the principal constituents to be those of kovar, with the presence of small amounts of copper and silicon explained by their adherence to the crust when the latter was removed from the cones and insulation.

The deposits were of unusual character. They were brittle, of low density, and of irregular structure. Larger particles 5 mils thick on the cones appeared to be formed of small, smooth globules less than 1 mil in diameter. Filamentary structures built of similar small particles were also present. In no way did the deposit give the appearance of sputtered material.

With sputtering of the cathode ruled out as the principal mechanism of the crust formation, another explanation was sought. The physical appearance of the cathode suggested one such explanation. This was that the cathode temperature exceeded the melting point of kovar, the molten surface then presenting a source of small droplets of metal, which by some means could be transferred to the surrounding areas, leaving behind the shiny, pitted cathode surface that was observed. To estimate the cathode temperature, it was assumed that cathode cooling was entirely by conduction, an assumption that proved reasonable since, at the cathode temperatures calculated, the heat loss by radiation was of the order of 1 percent of the total loss. Under this assumption, the tapered region of the cathode lead was treated as a conical portion of a hollow sphere (see fig. 1, b) of inner surface at temperature t_1, radius R_1; and outer surface at t_2, radius R_2. The fraction of the total surface of the sphere of radius R_2 intercepted by the tapered lead was equal to $A=2.4\times10^{-4}$ An approximate expression [2] for the heat conduction through the cone per second is then

$$h_1=\frac{4\pi K R_1 R_2}{R_2-R_1}(t_1-t_2)A. \qquad (1)$$

Here $K=0.046$ cal/cm (° C) sec is the coefficient of heat conductivity for kovar. Neglecting conduction through the glass walls, h_1 is the flow also through the cylindrical portion of length l, radius r, of the cathode rod to the external lead at $t_3=30°$ C, then

$$h_1=\frac{K\pi r^2}{l}(t_2-t_3). \qquad (2)$$

Now the discharge conditions were such that the maximum power dissipated at the cathode was no greater than 0.15 w. Taking then $h_1=0.15\ w=3.6\times10^{-2}$ cal/sec, eq 2 may be solved for the temperature t_2 of the junction of cylinder and cone. Equation 1 then yields a cathode temperature t_1 of approximately 500° C. This was 1,000 deg C below the melting point of kovar, precluding the melting of the cathode surface by this means.

Another and more promising explanation was that the abnormal glow might, due to some unknown surface conditions, break into an arc. Such an arc could be maintained for a short time despite the high impedance of the current source, if the tube leads, etc., provided sufficient distributed capacity to store the energy needed for the arc. Here would exist a means of locally heating the cathode, melting a small volume of the surface, and supplying the energy for its removal from the surface.

Heating of the cathode by this means has been discussed by Jurriaanse and Druyvesteyn [3] in a study of the breakdown from a glow to an arc discharge. They have determined that such arcing in a gas at a pressure of the order of 1 mm Hg requires a minimum energy storage of 2×10^{-7} j in the associated capacity. This makes necessary a capacity of the order of 6 $\mu\mu$f at the voltages encountered in the tube under study. These voltages (150 to 200 v) were essentially the same as those encountered by Jurriaanse and Druyvesteyn (150 v). Furthermore, arcing of this type was observed by these authors more frequently in argon than other noble gases; and oxidation of the cathode, quite likely to occur in the TR due to its water vapor content, was found to be essential to the breakdown.

With this in mind, distributed capacities of several tubes were measured with the tubes in place in the life test circuit, and typical values were found to be of the order of 20 $\mu\mu$f. This value is about three times the minimum reported by Jurriaanse and Druyvesteyn. Thus the voltage, capacity, and gas content of the modified 1B24 were certainly favorable for the production of arcs. It remained to be found if the energy in such arcs was ample for local melting of the cathode and removal of the molten drops, and also if the arcs could be observed experimentally.

Under normal operating conditions stored energies of the order of $\frac{1}{2}CV^2=4\times10^{-7}$ j$=10^{-7}$ cal existed in the distributed capacity. This energy was sufficient to raise to the melting point and fuse a mass M of the cathode material, found by the expression

$$1/2\ CV^2=Mh_s\ (t_{mp}-t_1)+Mh_f.$$

The specific heat (h_s) and heat of fusion (h_f) of kovar were not available. Taking $h_s=0.16$ cal/g deg C

[2] M. W. Fishenden and O. A. Saunders, The calculation of heat transmission (H. M. Stationery Office, London, 1932).

[3] T. Jurriaanse and M. J. Druyvesteyn, Physica **3**, 825 (1936).

and h_f=48 cal/g, which are the values for iron, the most abundant component of kovar, M, was found to be 5×10^{-10} g. This mass of iron would occupy a hemispherical volume of diameter $d_1=2\times10^{-4}$ inch, which was of the order of magnitude of the pits observed on the cathode surface.

The energy for removal of the molten metal was estimated by a calculation of the change in surface energy of the fluid (surface tension × area) as the surface area of the M grams of cathode material was increased during removal. The small molten mass was assumed to occupy a hemispherical region at the cathode surface having an exposed area $\pi(d_1/2)^2$. Upon removal, the mass was considered to assume a spherical shape of area $4\pi(0.794\ d_1/2)^2$, leaving behind a hemispherical pit whose exposed area was $2\pi(d_1/2)^2$. The factor 0.794 arises from the fact that the volumes of sphere and hemisphere were equal. A value of surface tension (S=1,500 dyne/cm) not likely to be exceeded by the cathode material was chosen. The energy required, $E=4\pi S\ [d_1^2/8+(0.794\ d_1/2)^2-(d_1/4)^2]$ ergs, was found to be 4×10^{-11} cal. This was quite negligible by comparison with the energy available in a single arc discharge. The energy to carry the mass from cathode to anode was even smaller, of the order of 10^{-15} cal.

Visual and oscilloscopic observation of the glow discharge in the tubes showed conclusively that such arcs did occur, the discharge voltage dropping during the readily visible arcs as was expected. Old tubes operated on direct current showed bursts of ten to twenty arcs, spaced at 5 minutes to an hour or more apart. The pulsed tube yielded arcs in groups of one or two per pulse for several pulses at irregular intervals. No statistical study of arc frequency was made, but it was established that the frequency was considerably higher than the average required to supply the deposit at a rate of M grams per arc.

IV. The Experimental Diodes

To learn more of the parameters affecting the arcing, diodes were constructed with which to determine whether certain factors were important in the formation of arcs or deposits and whether arcing was essential to the formation of the deposits. The factors were distributed capacity, cathode diameter and material, pulsing of the discharge current, water vapor in the gas fill, and shape of the anode. Electrodes of various types were used, but in all cases they were arranged as in figure 2. This represented a considerable simplification from the modified 1B24 construction, but preserved the characteristics of its construction such as cathode shape, area and insulation, and electrode spacing, essential to the problem. An advantage of the construction was that the inherent capacity in the tube was less than 1$\mu\mu$f. Additional capacity could be connected externally between cathode and anode in any amount desired.

As may be observed in figure 3, four different cathode types were used: 0.030 in. and 0.020 in. kovar glassed over except at the cathode face by about 0.030 in. of 7052 glass, 0.020 in. kovar, and 0.020 in. 18–8 stainless steel insulated by porcelain

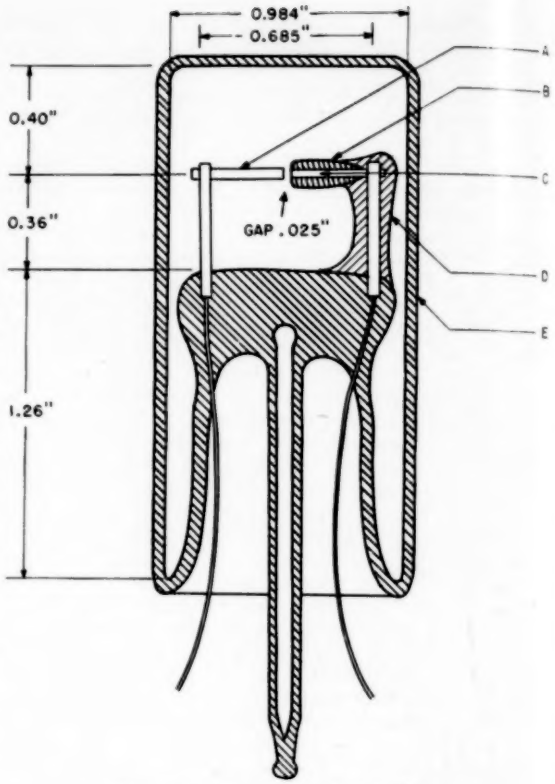

FIGURE 2. *Experimental diode number 2.*

A, anode 0.040-in. Ni rod; *B*, cathode insulation, 7052 glass; *C*, cathode, 0.030-in. kovar rod; *D*, support insulation, Sauereisen number 78; *E*, envelope and press of soft glass.

tubing. Anodes were all of grade A hydrogen fired nickel. These were 0.030 in. or 0.040 in. in diameter with either a flat face or a conical point similar to a 1B24 cone tip.

Table 1 shows the combinations in which the variables listed above were studied in relation to deposit rate. In the table, all variables within a group are held constant except those designated by bold-faced type. It is suggested that a relation between the varying parameters and the deposit resulting may be obtained readily by observing the bold-faced data and the corresponding brief description of the deposit in the right-hand column.

It may be noted that in each group there was a repetition of important combinations from the preceding group of tubes, which obviated concern over variation of conditions from group to group of tubes. Indeed this demonstration of consistent results from group to group indicated that comparison between the groups was permissible, yielding more extensive information than the bold-faced data suggest. However, the data printed in bold face is adequate for this discussion.

Ten of the tubes were operated under direct current conditions, while the remaining twenty were

TABLE 1. *Cathode deposits produced in the experimental diodes*

[Variables within a group were constant, except those printed in bold-faced type. For all tubes, the nominal run was 100 hr., electrode gap was 0.025 in., and maximum current was 0.75 ma. A relation between the varying parameters and the deposit produced may be obtained by observing the data in bold-faced type, and the corresponding description of the deposit in the right-hand column]

Group	Tube No.	Cathode			Anode (Ni)		Gas fill	Capacity	Current type	Deposit (brief description)
		Material	Insulation	Diameter	Shape of face	Diameter				
				in.		*in.*		*μμf*		
1	2, 8	Kovar	Glass	0.030	Flat	0.040	A+H_2O	0	d-c	Negligible.
	9	do	do	.030	do	.040	A+H_2O	570	d-c	Very small.
	4, 5	do	do	.030	do	.040	A+H_2O	**1,100**	d-c	Small.
2	11	do	do	.020	do	.030	A+H_2O	0	d-c	Negligible.
	14	do	do	.020	do	.030	A+H_2O	570	d-c	Small.
	12	do	do	.020	do	.030	A+H_2O	**1,150**	d-c	Moderate.
3	16, 19	do	do	.020	do	.030	A+H_2O	0	**Pulse**	Negligible.
	20	do	do	.020	do	.030	A+H_2O	0	**d-c**	Do.
	17, 18	do	do	.020	do	.030	A+H_2O	**1,100**	**Pulse**	Very large.
	21	do	do	.020	do	.030	A+H_2O	**1,100**	**d-c**	Moderate.
4	24, 25	do	do	**.020**	Pointed	.040	A+H_2O	1,100	Pulse	Very large.
	22, 23	do	do	**.030**	**Pointed**	.040	A+H_2O	1,100	do	Large.
	26, 27	do	do	.030	**Flat**	.040	A+H_2O	1,100	do	Do.
5	28, 30, 32	**18-8 steel**	Ceramic	.020	Pointed	.040	A+H_2O	1,100	do	Negligible, except 32, small.
	29, 31, 33	**Kovar**	**Ceramic**	.020	do	.040	A+H_2O	1,100	do	Largest obtained.
	35	Kovar	**Glass**	.020	do	.040	**A+H_2O**	1,050	do	Very large.
	36, 37	do	Glass	.020	do	.040	**A only**	1,050	do	None.
	34	do	do	.030	Flat	.040	A+H_2O	1,070	do	Large.

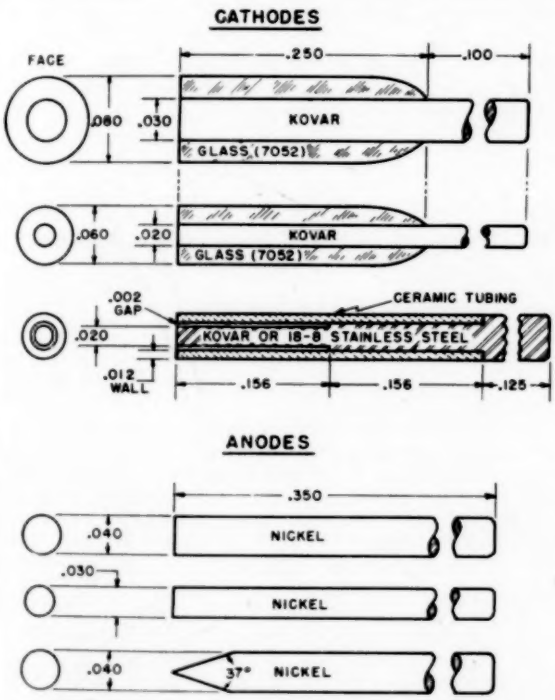

FIGURE 3. *Electrode types used in experimental diodes.*

Dimensions are in inches.

pulsed, using a thyratron circuit to provide a quarter sine wave of current with fast rise and sinusoidal decay at a 60 c/s repetition rate.

During the arc, extra current was drawn momentarily from the pulser due to reduced tube resistance. This produced a pip super-imposed on the quarter wave trace obtained by connecting an oscilloscope across a resistor in series with the tube, thus indicating the occurrence of the arc. The arc rate and capacitor energy were found to be in proportion to the deposit in cases where these factors were measured.

The gas fill was the same as that in the modified 1B24 with the exception of the series in which water vapor was omitted to study its effect upon the deposit. Each of the tubes was run approximately 100 hr.

It was established that arcing occurred in the diodes at the capacity measured for the 1B24. Larger values were used in the runs, however, to increase the size of the deposits and to make their comparison easier, as deposit measurement was accomplished by tracing enlarged shadowgraphs of the electrodes seen on a microcomparator before and after the run.

Figure 4 presents sample tracings obtained from the shadowgraphs before and after 100-hr runs. On these may be seen the irregular conductive deposits, which, as the cathode deteriorates, build up on the insulator surrounding the cathode to provide the shorting bridge.

V. Discussion of Diode Results

Early experiments established that the deposits observed in the diodes were proportional to the arc rates in these tubes. Thereafter, the deposit shadowgraph made at the end of a 100-hr run was used as an indication of the total number of arcs that had occurred in a given tube. An occasional check on arc rate, however, was maintained visually during the runs. By study of a tabulation of the varying parameters similar to table 1, together with a complete series of shadowgraphs for the tubes, it was possible to determine the principal factors in the formation of the short circuiting deposits.

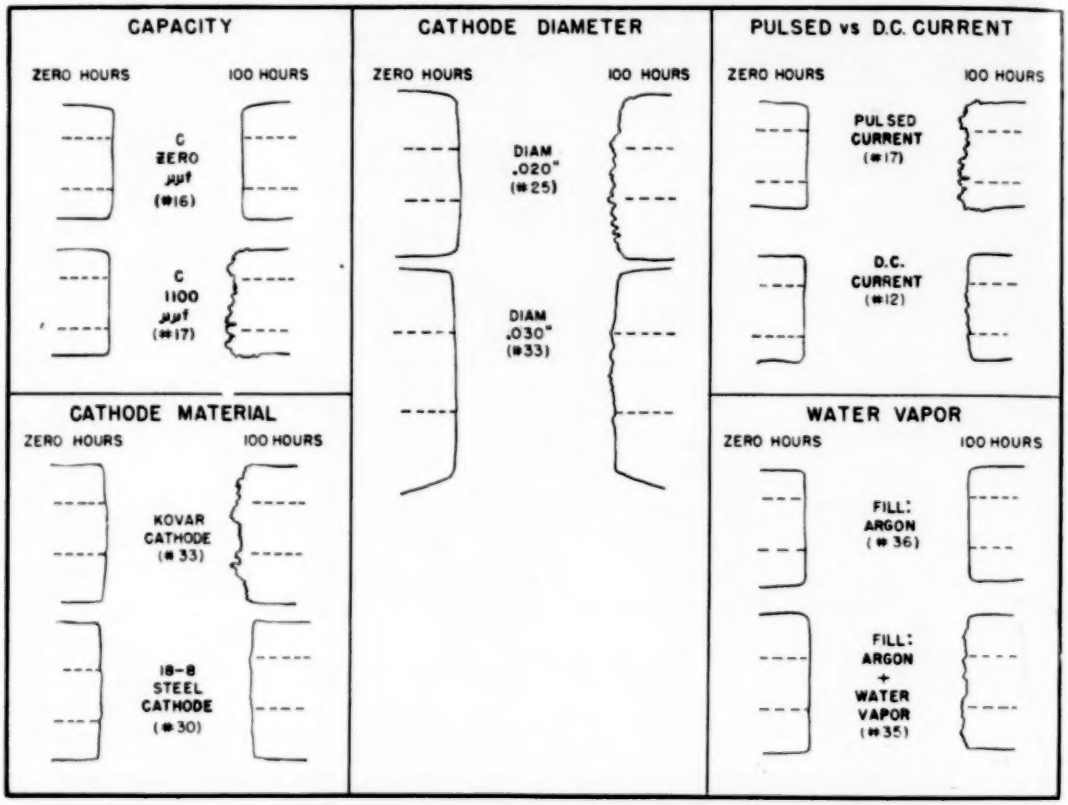

FIGURE 4. *Examples of cathode tracings.*

Microcomparator tracings of cathode profiles at the beginning and end of a 100-hr run show the effects produced on the deposit by variation of five tube parameters Dotted lines indicate the location of the cathode conductor within its insulation.

1. In the absence of external capacity when the total capacity at the tube of electrodes and circuit wiring is of the order of 1 $\mu\mu$f, neither arcs nor deposits of the type under consideration occurred. The capacity in this case was but one-twentieth of that of the 1B24 in its life test circuit. Upon addition of external capacity both arcs and deposits appeared, and the magnitude of the deposit increased for larger capacity. The arc as the mechanism of the deposit and the necessity for arc energy storage in the distributed capacity were clearly established.

2. Decreasing cathode diameter with attendant rise in current density at the cathode resulted in an increasing deposit rate. The arcing rate increased at the same time.

3. The anode shape was not found to be a particularly important factor. The fact that the pointed anode had little effect on the anode deposits indicated that the deposited substance was uncharged.

4. Pulsing of the discharge increased the deposit rate over that obtained under direct-current conditions. It is probable that surface conditions necessary for the arc breakdown had arisen between current pulses in this case, while a direct-current discharge may have discouraged these conditions by constant cleaning of the cathode.

5. Water vapor played a decisive part in the formation of the deposits. Neither arcs nor deposits were obtained in the tubes from which water vapor was omitted, indicating that this vapor may be essential to establishment of the surface conditions mentioned before. This fact, together with (1) and (4), is in accord with the findings of Jurriaanse and Druyvesteyn previously mentioned.

6. A most important factor is the cathode material. Eighteen-eight stainless steel represents a great improvement over kovar. As shown in the table previously, for only one of three 18-8 cathodes were arcs observed at all. This arcing was delayed for 60 hr, and the small deposit that was formed was highly localized by comparison with those on the kovar electrodes. The remaining 18-8 cathodes showed negligible deposits. Sample tracings of kovar and 18-8 cathodes (fig. 4) clearly show the relative extent of deposit.

Since glass does not seal to 18-8 steel, it was necessary to use a slip-on insulator, in this case porcelain, with this metal (fig. 3). The electrode design

provided a narrow gap between the cathode and the ceramic to avoid a probable one point contact between cathode and ceramic in the absence of a perfect fit. Because all of the kovar cathode tests had been run with sealed-on glass insulation, a control test was made in which the relative deposit rates of glass- and porcelain-insulated kovar cathodes were measured (table 1, group V, tubes No. 29, 31, 33, 35). This control test showed a somewhat larger arc rate in the latter case, which at first thought appears contrary to the practice, generally accepted in tube development, of discouraging arc back by just such separation of a negative electrode from its insulation. The practice arose after the study by Issendorf, Schenkel, and Seeliger [4] of the effect on arc back of various insulating materials in contact with a negative electrode. In the present case, the explanation of the apparent discrepancy probably lies in the difference of insulator material, and use of quartz, alundum, or even glass instead of porcelain might have avoided the discrepancy. The question was not pursued, since the improvement derived from use of 18-8 stainless steel cathodes far outweighed the disadvantage of the insulator choice, leaving no question as to the importance of the choice of cathode material.

The remedial steps suggested by this work are clear. Distributed capacity should be reduced as far as possible. Pulsing the discharge should be avoided when permissible, and the cathode current density should be as low as possible. Ideally, water vapor or other active gas or vapor should be omitted from the tube. This is an unobtainable ideal in most cases; an electronegative component is usually required in a TR tube in order to keep recovery time adequately short. The steps mentioned so far are generally limited in their feasibility by the particular application of the tube. One factor, the cathode material is usually not so determined. It has been shown that selection of this material is a critical factor in determining the life of a tube. It should therefore be possible to extend the life of TR tubes operating under the assumed conditions, where life is prematurely ended by keep-alive short circuit, by substitution of a ceramically insulated 18–8 stainless steel cathode for the customary glass insulated kovar cathode. The results described here might well be of value in the design of any electron tubes wherein high current density glow discharges are maintained between electrodes in close proximity.

The generous and valuable advice of John E. White, Chief of the Bureau's Electron Tube Laboratory, and of Wm. G. Dow of the University of Michigan is gratefully acknowledged.

WASHINGTON, May 11, 1950.

[4] J. von Issendorff, M. Schenkel, and R. Seeliger, Wiss. Veröff. Siemens-Werk. IX, 1, 73 (1930).

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2140

Test Mixtures for Distillation at Atmospheric and Reduced Pressure[1]

By Charles B. Willingham[2] and Vincent A. Sedlak[3]

In this paper are reported values for the separation factor for distillation for each of the three pairs of diethylbenzenes for pressures from 30 to 800 mm Hg. Also given are data on the refractive index as a function of composition for the system 1,2-diethylbenzene and 1,3-diethylbenzene.

A mixture of 1,2-diethylbenzene and 1,3-diethylbenzene was selected on the basis of boiling point, separation factor, and expected ideality for use as a test mixture for investigating the separating efficiency, at pressures in the range from 30 to 760 mm Hg, of the rotary concentric-tube [1,2][4] and other distilling columns [3]. As there is a difference of 0.0078 in the values of the refractive index, n_D at 25° C, of the two pure components, it is possible to analyze this mixture refractometrically within one and one-half percent of the composition from measurements of refractive index made to the nearest 0.0001.

Accordingly, known mixtures of 1,2-diethylbenzene and 1,3-diethylbenzene, having compositions near 25, 50, and 75 percent of the former, were made up accurately by weight, by using samples of purified compounds corresponding to the material prepared for standard samples [4]. Measurements of the differences in the refractive index of these mixtures and of pure 1,2-diethylbenzene, with respect to pure 1,3-diethylbenzene, were made by using a modified Abbe-type precision refractometer [5]. The results, expressed as the difference between the refractive index of the mixture and of 1,3-diethylbenzene, are as follows:

For 1,3-diethylbenzene and 1,2-diethylbenzene, at 25° C

$$n_D(\text{mixture}) - n_D(1,3\text{-diethylbenzene}) = 0.00782N - 0.00008N^2,$$

[1] This investigation was performed at the National Bureau of Standards as part of the American Petroleum Institute Research Project 6 on the Analysis, Purification, and Properties of Hydrocarbons.

[2] Now with the Mellon Institute, Pittsburgh, Pa.; formerly with the National Bureau of Standards.

[3] Now with the American Petroleum Institute Research Project 6 at the Carnegie Institute of Technology; formerly a Research Associate at the National Bureau of Standards.

[4] Figures in brackets indicate the literature references at the end of this paper.

where n_D is the refractive index, and N is the mole fraction of 1,2-diethylbenzene in the mixture. The foregoing equation reproduces the experimentally observed values of refractive index within ± 0.0001.

The number of equivalent theoretical plates at total reflux in a rectifying section may be calculated from the equation [6,7]

$$n+1=[1/\log\alpha]\ \log\ [(N_A/N_B)^{\text{head}}/(N_A/N_B)^{\text{pot}}],$$

where

$$\alpha=P^*_A/P^*_B.$$

P^* is the vapor pressure of the given component in the pure state at the mean temperature, N is the mole fraction of the given component in the liquid phase in the head or pot as indicated, and A and B represent the more volatile and the less volatile component, respectively, in an ideal binary solution.

Binary mixtures of the isomeric diethylbenzenes are expected to be substantially ideal as regards the relation of vapor pressure to composition (high efficiency distillations of aromatic concentrates have disclosed no nonideality between close-boiling isomeric aromatic hydrocarbons) [8,9]. The data on the vapor pressure as a function of temperature of the pure components recently reported from this laboratory [10] were reduced for each of the three pairs of isomers to give values of α at mean temperatures corresponding to average vapor pressures between 30 and 800 mm Hg. For convenience in evaluating the separating power of distilling columns, the separation factor, $1/\log_{10}\alpha$, as a function of the logarithm of the average pressure for the components, between 30 and 800 mm Hg is presented graphically in figures 1 and 2.

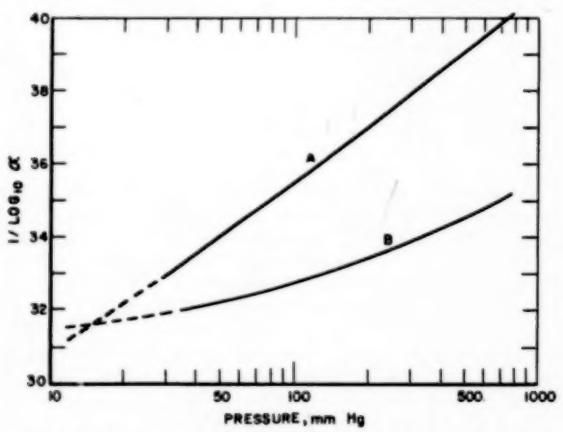

FIGURE 1. *Values of the separation factor as a function of the logarithm of the average pressure.*

The scale of ordinates gives the values of the separation factor, $1/\log_{10}\alpha$, and the scale of abscissas gives the values of the logarithm of the average pressure in mm Hg. Curve A is for the mixture of 1,3-diethylbenzene and 1,2-diethylbenzene. Curve B is for the mixture of 1,3-diethylbenzene and 1,4-diethylbenzene.

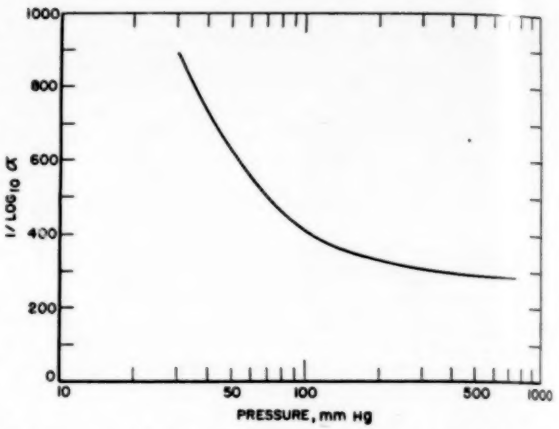

FIGURE 2. *Values of the separation factor as a function of the logarithm of the average pressure.*

The scale of ordinates gives the values of the separation factor, $1/\log_{10}\alpha$, and the scale of abscissas gives the values of the logarithm of the average pressure in mm Hg for the mixture of 1,2-diethylbenzene and 1,4-diethylbenzene.

At 1 atm pressure, the values of the boiling points for the three isomers are as follows [10]:

1,4-diethylbenzene, 183.75° C;

1,2-diethylbenzene, 183.42° C;

1,3-diethylbenzene, 181.10° C.

The values of α at the mean boiling point at 1 atm for each of the three pairs of isomers are calculated from the data of reference [10] as follows:

1,3-diethylbenzene and 1,4-diethylbenzene, 1.0677;

1,2-diethylbenzene and 1,3-diethylbenzene, 1.0591;

1,2-diethylbenzene and 1,4-diethylbenzene, 1.0081.

Although only one of the three pairs of isomers may be analyzed as a mixture refractometrically, all three pairs may be analyzed by infrared spectrophotometry [11], since purified standard samples of each of the three diethylbenzenes are now available [4] for spectrometer calibration.

Grateful acknowledgment is made to Frederick D. Rossini for his suggestions in the course of this investigation, and to Alphonse F. Forziati for the measurements of refractive index.

[1] C. B. Willingham, V. A. Sedlak, J. W. Westhaver, and F. D. Rossini, Ind. Eng. Chem. **39,** 706 (1947).

[2] C. B. Willingham, V. A. Sedlak, J. W. Westhaver, and F. D. Rossini, Test of the rotary concentric-tube distilling column at reduced pressure.

[3] C. B. Willingham and F. D. Rossini, J. Research NBS **37,** 15 (1946) RP1724.
[4] A. J. Streiff, J. C. Zimmerman, L. F. Soule, M. T. Butt, V. A. Sedlak, C. B. Willingham, and F. D. Rossini, J. Research NBS **41,** 323 (1948) RP1929.
[5] A. F. Forziati, J. Research NBS **44,** 373 (1950) RP2085.
[6] M. R. Fenske, Ind. Eng. Chem. **24,** 482 (1932).
[7] M. R. Fenske, Science of Petroleum, **II,** 1629. (Oxford Univ. Press, London, 1938).
[8] A. F. Forziati, C. B. Willingham, B. J. Mair, and F. D. Rossini, J. Research NBS **32,** 11 (1944) RP1571.
[9] A. F. Forziati and F. D. Rossini, J. Research NBS **39,** 425 (1947) RP1839.
[10] A. F. Forziati, W. R. Norris, and F. D. Rossini, J. Research NBS **43,** 555 (1949) RP2049.
[11] American Petroleum Institute Research Project 44, National Bureau of Standards. Catalog of Infrared Spectral Data, Serial No. 326, 1,2-diethylbenzene, contributed by the Socony Vacuum Oil Co., and Serial Nos. 331, 332, 333 for the three isomers of diethylbenzene, contributed by the Anglo-Iranian Oil Co.

WASHINGTON, June 1, 1950.

ournal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2141

Aging of Karakul and Seal Fur Skins

By Edward T. Steiner and Elizabeth R. Hosterman

Physical and chemical data are presented for two types of fur skins of widely different dressings and characteristics to show that deterioration takes place when furs are stored under conditions considered noninjurious. There occur during storage an appreciable loss of strength amounting to as much as 25 to 40 percent, and a change in certain chemical properties. Pelts weakened by aging showed increased soluble nitrogen and water extractives and lower grease contents (petroleum ether extractives) than less aged skins. A fair correlation exists between the observed changes in chemical and physical properties. The presence and quantity of copper in the weakened pelts offer the only reasonable explanation for the cause and the extent of the deterioration.

I. Introduction

The value of dressed fur skins used annually in his country is estimated at several hundred million lollars. The more than moderate investment that s required in the purchase of fur garments does not lways seem to be commensurate with the expected - to 10-year service life of most furs. By contrast, roperly tanned leather may be expected to last rom 10 to 40 years when not subjected to severe brasion or chemical action. If the causes for the elatively short life span of the processed furs could e discovered and eliminated, there would result onsiderable economic advantage to the consumer, he distributor, and the manufacturer.

Various reasons have been offered for the relatively oor wearing qualities of fur garments. In the first lace, the skins from most fur animals are naturally eak. Secondly, the procedures for dressing the kins are not so highly developed as those for making eather, primarily because of the purposes for which he furs are intended. In straight leather making, emoval of the hair is required. The process of emoving it tends to give better quality leather. With furs, the hair, of course, must be left on. Cerain chemicals and tanning agents suitable for leather nanufacture adversely affect fur quality and appearance, and cannot be used for fur dressing. It is the im of fur dressers to keep the weight of the dressed elts light, while the leather manufacturer usually trives to give weight to his product.

Carelessness on the part of the consumer is often esponsible for the poor service given by furs. Improper drying after exposure to the elements, faulty leaning methods, and storage in polluted atmospheres materially shorten the life of fabricated furs. Such factors, however, do not account for the loss of strength of furs that have never been exposed to such conditions. Fur skins frequently become weak with age, apparently without cause.

The source of the deterioration is not apparent because the leather is concealed from view between the hair of the fur and the lining of the garment. The present work was undertaken to investigate the possible causes and extent of deterioration occurring in fur skins stored under conditions that would not be expected to cause losses of strength.

The Government's interest in furs centers about a protective and advisory role. It wishes to assist and cooperate in helping to bring about maintenance of a constant supply of fur animals, whether they be wild or captive. Through its agency, the Fish and Wildlife Service, Department of Interior, the Government compiles results of surveys of state agencies on the stock of various species and enacts legislation based on these statistics, regarding the "take" of the animals, in order to provide a balance between supply and preservation. Through another of its agencies, the Fur Farming Investigations, of the Department of Agriculture, economical methods of feeding, breeding, and management of fur animals in captivity are studied.

Examples of the work the Government has been doing along these lines include supervision of the annual catch of Alaskan fur seals off the Pribilof Islands and experimentation in the breeding of Karakul sheep.

Alaskan seal furs are furs of distinction and importance in the industry. By far the world's largest supply comes from the Pribilof Islands off the coast of Alaska. Until 1910 pelagic sealing, which means the killing of seals in water, almost resulted in the

extinction of the herd. Through the efforts of this country, a treaty that prohibited the pelagic sealing operations was drawn up between the United States and other sealing countries. Fur sealing operations are now directly under the control of the Federal Government, and as a result the herd has grown from a state of near depletion to about 3,000,000. Each year the Government, after estimating the number of animals, specifies how many seals may be taken from the herd for fur purposes. The killing is made under Federal supervision, after which the skins are shipped inland, processed under Government contract, and sold at public auction. The income from the seal industry to date has netted the United States approximately $100,000,000.

Karakul type lambs are valuable in the manufacture of fur coats and fur-trimmed garments. Since importation of the live animals from countries in which the breed is being grown is prohibited by legal restrictions, because of serious diseases of the animals, multiplication in this country must be through breeding and cross-mating of Karakuls.

Since 1928 the Bureau of Animal Industry, U. S. Department of Agriculture, has been conducting cross-breeding experiments in which Karakul rams have been mated with ewes of the more available American types. The purpose of the cross-breeding is to compare the quality of fur produced thereby with fur produced on purebred Karakuls. The value of the yearly crops of the resultant skins is judged for that Bureau by appraisers in the Persian lamb fur industry, after which the skins are dressed according to commercial procedures.

II. Description of the Furs

The Alaskan seal furs were used as standards of comparison in 1939 in evaluating tannages of various bidders in connection with contract awards for the dressing of the Government's annual take of seals from the Pribilof Islands. Replicate samples of the standards were stored under conditions generally considered satisfactory for preservation, in the National Bureau of Standards leather laboratory, for 10 years. A limited number of properties were studied in 1939; the same tests were repeated on the storage samples in 1949. Some of the results were found suitable for comparison with those obtained on the lambskins.

Case histories of the seal furs are not known. Each pelt appeared to be of good quality, the skin being dressed by an oil process, and the fur dyed either an attractive brown or black shade.

A limited number of the dressed lambskins were made available to this Bureau through the courtesy of the Bureau of Animal Industry, as a part of a cooperative project relative to investigations on the effect of dressing procedures and other factors on the strength and durability of fur skins.

The Karakul pelts are excellent subjects for investigation, because of the wide age range of from 3 to 20 years. Unfortunately, information concerning the freshly processed skins is not available. In studying aging effect, certain physical and chemical

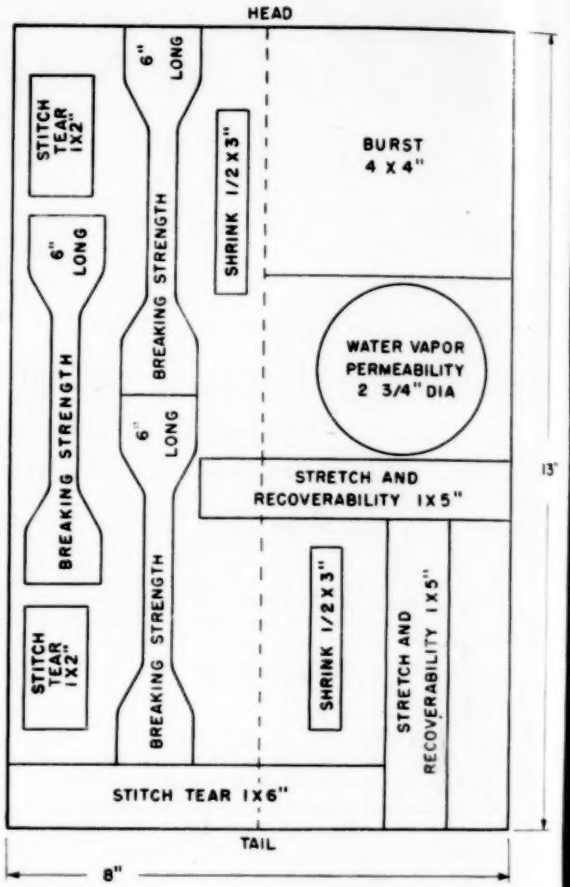

FIGURE 1. *Location of test pieces for physical tests on lambskins.*

properties of the fur skins can be defined that are indicative of the changes that take place as the storage period of the skins increases.

Specimens of the lambskins, representative of the yearly crops of 1929, 1930, 1931, 1932, 1933, 1944, 1945, and 1946, were submitted. The number of skins available did not permit sampling an equal number for each year group. In addition, pelts within a year were composed of purebreds and crossbreds, and it was necessary to observe whether aging qualities of purebred Karakuls differed from those that were crossbred. Specimens of each group were selected for physical tests in accordance with the scheme shown in figure 1. Remnants from the physical test pieces, as well as the remainder of the pelts from which they were cut, were ground to serve as samples for chemical tests.

III. Effect of Crossbreeding in Relation to Strength of Karakul Pelts

A comparative study of the effect of breeding was made on three skins of 100-percent Karakul, and on three skins of 50-percent Karakul, for the years 1930,

1931, and 1932. Breaking strength and stitch tear A values were determined in accordance with the methods outlined in Federal Specification KK–L–311, Leather and Leather Products; General Specifications (Methods of Sampling, Inspection, and Tests).

Stitch tear B specimens were prepared in the following manner. The two free ends of the 1- by 5-in. specimens were sewn together by a local furrier who used equipment and technique of the industry. The resultant loop was then cut with scissors so that the stitched seam was in the center of the specimen. The ends were then placed in the jaws of a tensile testing machine, and the force required to pull the stitchings through the leather was measured.

Burst values were obtained on equipment modified for testing leather by Bureau personnel. The apparatus fastens to the cage used in compression measurements on a machine of the Tinius Olsen Electromatic Tester type. The principle is essentially the same as for the usual Mullen burst tester, except that the plunger is entirely metal and of smaller diameter than the Mullen's rubber diaphragm. The modified equipment has greater testing capacity than the Mullen and is capable of testing leathers of high stretch values. The plunger was forced against the flesh side of the leather.

The average values for the physical properties of the two breeds are shown in table 1, for the years specified.

TABLE 1. *Comparison of strength and purity of breed of aged Karakul pelts*

Test	Year of crop			Purity of breed
	1930	1931	1932	
				Percent
Breaking strength (lb/½-in.-wide strip)	22.3	26.4	23.0	100
	15.6	20.5	12.6	50
Stitch tear A (lb)	5.9	8.6	6.8	100
	4.4	5.9	4.6	50
Stitch tear B (lb)	3.9	9.4	6.0	100
	5.2	6.0	3.4	50
Burst (lb)	88.0	105.0	112.0	100
	65.0	86.0	53.0	50

The results show that, on the average, the difference of each physical property between the two purities of breed is greater than that which occurs or can be accounted for in the measurement of the individual specimens within a breed.

The data show highly significant differences due to breed at the 1-percent level for the first two properties listed, and differences significant at the 5-percent level for the stitch tear B and burst values. Significance at the 1-percent level means that if no real difference existed, then a difference as large as that observed or larger would occur only once in a hundred times on the average. Significance at the 5-percent level means that if no real difference existed, then a difference as large as that observed, or larger, would occur only five times in a hundred times on the average. Therefore, one can assume that a real difference exists between the two breed types of sheep and that the degree of confidence one can have in the difference, as represented by the results of the various tests, is greater for the tests of 1-percent significance than those of 5-percent significance.

In finding the aging effect within a breed, only the 100-percent Karakuls were studied. Sufficient specimens of the 50-percent crossbred type were not available for similar comparison.

IV. Physical and Chemical Properties of the Furs

The physical properties of the 100-percent Karakul skins for various years of age are shown in table 2.

All of the determinations were made on test pieces with the fur on, as it is practically impossible to remove the fur without damaging the test pieces. Efforts to remove the hair by means of clippers, electric razors, buffing, and the like were unsuccessful, and the results of the strength data are expressed, therefore, without regard to thickness.

There is no evidence of yearly or gradual changes in physical properties, particularly strength, during the storage periods. However, if the skins of 1930, 1931, 1932, and 1933 are grouped and compared, with the skins of 1945 and 1946 considered as another group, significant differences are indicated insofar as breaking strength, furrier sewn-stitch tear (B), and burst values are concerned.

TABLE 2. *Physical properties of aged 100 percent Karakul fur skins*

Property	Year of crop					
	Group I				Group II	
	1930	1931	1932	1933	1945	1946
Breaking strength (lb)	22.3	26.4	23.0	19.3	38.2	33.5
Stretch at break (lb)	31.3	37.0	31.2	35.9	37.2	32.4
Stitch tear A (lb)	5.9	8.6	6.8	7.4	6.6	6.2
Stitch tear B (lb)	3.9	9.4	6.0	9.4	13.5	14.1
Burst (lb)	88.0	105.3	112.0	89.0	205.3	142.0
Water vapor permeability (mg/min)	4.5	4.9	5.0	5.4	4.6	4.5
Resilience coefficient (parallel to backbone)	0.24	0.26	0.22	0.26	0.24	0.21
Shrink temperature (° C)	55.8	53.7	48.7	49.0	52.7	54.0

A comparison of data in groups I and II indicates that there has been loss in strength of the lambskins during the longer storage periods.

On the assumption that the loss of strength is caused by a decomposition process, the six age groups were examined chemically. The constituents of the skins determined were moisture, grease (petroleum ether extract), total and insoluble ash, water solubles, and total and soluble nitrogen. The results are tabulated in table 3, on a dry combined leather and hair basis. The same difficulty in removing the hair from the chemical test specimens was experienced as with the physical test pieces. The incompleteness with which the fibers could be removed from the grain, and the uncertainty of how much remained within the grain after clipping or shearing, prompted the decision to sample the whole pelt for the comparative chemical analysis.

904428—50——5

TABLE 3. *Chemical constituents of the yearly crops—100-percent Karakul*

Constituent	Year of crop					
	Group I				Group II	
	1930	1931	1932	1933	1945	1946
	Percent	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*
Grease (dry)*	6.97	6.27	9.65	8.41	10.55	10.27
Ash, total (dry)	5.60	6.64	6.70	4.07	5.52	5.51
Ash, insoluble (dry)	2.10	2.79	2.20	2.02	2.31	1.98
Water solubles (dry)	12.12	9.89	10.34	8.92	5.60	6.53
Protein, total (dry)	75.42	74.80	68.73	72.78	69.74	71.15
Protein, soluble (dry)	4.67	3.15	4.10	3.48	1.24	1.18
pH	3.5	3.5	3.5	3.4	3.3	3.4

* Petroleum ether extract.

Although the percentage of hair or wool may vary from pelt to pelt, the procedure of including the hair in the examinations is probably no less accurate than one using samples from which the hair had been only partially removed. The factor 5.62 is used to convert the nitrogen value to protein.

As in the data for the physical properties, there seems to be no yearly or gradual change in chemical properties. If, however, the skins are again grouped and compared as follows: 1930, 1931, 1932, 1933 versus 1945 and 1946, significant differences are noted in the grease (petroleum ether extract), water solubles, and soluble protein of those two groups.

The amounts of grease extracted with petroleum ether are lower for the more aged skins. It is possible that some of the grease in the skins has volatilized during the long storage periods, but it is also possible that some free oils in the pelts have oxidized during the years of aging. The oxidized oils would not be recoverable by the petrolic solvent and would therefore appear as apparent losses of grease. Compared with group II, group I pelts show an approximately 25 percent lower petroleum ether extractive value. Using the same solvent, black and brown seal furs showed, respectively, reductions of 11.5 and 14.1 percent of their original grease contents.

In contrast to the "lower" grease content in the lambskins, the water soluble contents are higher for the more aged skins. This is to be expected, for if a deterioration is occurring it should be reflected in a higher proportion of water solubles, as soluble nitrogenous decomposition products are found in the water solubles.

Summation, on a dry basis, of the total protein value, grease, insoluble ash, and water solubles accounts fairly well for the total composition of the skins. The composition accounted for by the addition of these items, for the year groups, is as follows:

Year	Percentage	Year	Percentage
1930	96.6	1933	92.1
1931	93.7	1945	88.2
1932	90.2	1946	89.9

One-hundred-percent recovery is not expected since these skins are predominantly alum-tanned. The methods of analysis provide no means of accounting for all the components of the skins. It appears that as deterioration proceeds there occurs a breakdown of the mineral complex, which is reflected in the water solubles value. High water soluble values relate well with low residual strengths remaining in the skins. Not only are soluble nitrogenous decomposition products found in the water solubles, but other organic matter as well. If the water soluble values are corrected for soluble ash, and the remainder considered as organic matter, the latter correlate inversely with strength values. In table 4 are shown the calculated values of the residual organic matter compared with burst values.

TABLE 4. *Relation of soluble organic matter to burst values*

Year	Organic matter (calculated)	Burst	Year	Organic matter (calculated)	Burst
	Percent			*Percent*	
1930	8.62	88	1933	6.87	89
1931	6.04	105	1945	2.39	205
1932	5.84	112	1946	3.00	142

The correlation may be fortuitous, but these two properties certainly appear worthy of inclusion in observing changes occurring in aging fur skins.

The results of pH determinations on the lambskins preclude the possibility that the skins deteriorated because of red rot. The pH values lie in a range considered safe for leather preservation. Similarly, pH values of 3.5 and 3.0, obtained respectively for the black and brown seal furs after 10 years storage show that they are not dangerously acid.

Sulfur values, expressed for simplicity as $BaSO_4$, for the two age groups of the lambskins are:

	1930 to 1933	*1945 and 1946*
	%	%
Total	12.3	12.2
Soluble	1.9	1.9

The values are high for both groups because of the inclusion of hair in the sample. The data show no change in sulfur content, and consequently no increased sulfuric acid content. It seems safe to assume that no acid deterioration has taken place in the skins.

V. Effect of Copper and Iron on the Strength of the Leather

Qualitative tests of the ash of the six groups of lamb furs showed the principal constituents, other than alkali salts, to be aluminum, iron, copper, and some phosphates. Aluminum probably constitutes the tanning agent, while the copper and iron are integral parts of the dyeing process.

Determinations of the amounts of these metals in the ash show that the ratio of each has not changed appreciably for dressings of each year. It would

appear from the totals that the tanning and dyeing of the furs from year to year had not been changed to the extent of altering the characteristics of the pelts. The amounts found are shown in table 5.

TABLE 5. *Metallic constituents of ash (expressed in percentages of leather plus hair)*

Constituent	Year of crop					
	Group I				Group II	
	1930	1931	1932	1933	1945	1946
	%	%	%	%	%	%
Al_2O_3	1.06	1.14	0.82	0.67	0.89	0.81
Fe_2O_3	0.54	0.63	.57	.69	.65	.78
CuO	.30	.22	.32	.34	.14	.18
Total	1.90	1.99	1.71	1.70	1.68	1.77

The copper values differ significantly (at the 5-percent level) between the two age groups. This may be meaningful. It has been postulated by a number of investigators that when copper and iron are present as impurities in vegetable-tanned leathers, deterioration is accelerated. Kanagy,[1] in noting the lack of agreement in the conclusions of the investigators, conducted accelerated aging tests by exposing quebracho and chestnut-tanned leather impregnated with various salts of copper and iron, to an atmosphere of oxygen at 100 lb/in.^{2} and 100° C. His conclusions, in part, are as follows: All leathers containing added copper showed more deterioration than corresponding untreated leathers. Most of the deterioration occurs above pH 3. The increase in the rate of deterioration with increasing copper content is most rapid for amounts less than 0.2 percent. At pH values above 3.5, the effect of iron on the rate of deterioration is not appreciable. At lower pH values the deterioration appears to be a function of pH rather than iron content. For equal amounts of the metal added, that in the form of the sulfate was more active than that added as the acetate or lactate.

It is to be noted that the copper content for the strongest skins is less than that for the weaker skins. The content correlates fairly well with the yearly degradations.

The only physical test, as an indication of strength, made on the seal samples in 1939 was breaking strength. This property remeasured 10 years later on replicate samples was considerably lower, indicating deterioration during the storage period for both the brown and the black-dyed samples. Black-dyed pelts had lost more strength than brown-dyed pelts. Prior to storage, the pelts of both colors had approximately the same strength. Qualitative chemical analyses showed that the black dye was produced mainly with iron and copper salts, while the brown was achieved by the use of salts of tin, copper, chromium, and iron. Other metals were present in very small amounts in both varieties, but the magnitude of the amount did not justify quantitative analysis.

Comparison of available data, both physical and chemical, of the Karakul and seal furs are shown in table 6.

TABLE 6. *Comparison of Karakul and seal furs*

Test	Condition	Fur		
		Black seal	Brown seal	Karakul
Breaking strength (lb)	Original	56.0	58.0	35.9
	Aged	37.8	42.0	22.8
Breaking strength loss (%)	Original			
	Aged	38	28	36
CuO (%)	Original	(a)	(a)	0.16
	Aged	0.41	0.28	.30
Fe_2O_3 (%)	Original	(a)	(a)	.72
	Aged	0.63	0.07	.61

a Not determined.

The terms "original" and "aged" are used to describe the relative age levels of the furs. The original Karakuls were in storage approximately 5 years, while the original seal fur age was less than 1 year.

The three types of furs lost between 30 and 40 percent of their original strength after storage periods of approximately 10 years. Both Karakul and black seal furs contained considerable quantities of iron. The iron values of group I and group II of the lamb furs do not vary significantly. The quantity of iron found in the brown-dyed seal furs can hardly be considered the cause for the extent of deterioration that has occurred. The iron salts used in the dyeing, therefore, do not appear to be responsible in causing the breakdown of the pelts.

In each instance of comparison, higher copper values are associated with the greater losses of strength. It is concluded then that copper is instrumental in causing the deterioration. The quantities of the metal found in the skins suggest that its action is catalytic.

The deductions concerning the cause for the deterioration of the furs are in good agreement with the conclusions advanced by Kanagy, although the conditions of test and the materials used are quite different.

Kanagy's research involved accelerated aging by heat and pressure in the laboratory, while this study deals with naturally aged skins; his experiments were on vegetable-tanned heavy leather, quite unlike the light alum-and-oil-tanned skins studied in the present work; and copper was added to his leathers for the purpose of the investigation instead of being used as a normal ingredient in the dressing process, as was the case in this study.

So far as is known, this study is the first in which catalytic degradation of oil- or alum-tanned skins by copper has been indicated.

Since the copper percentages of the seal skins are higher than the percentages in the lambskins, a greater extent of deterioration might also have been expected. This need not necessarily be so, as the dye coloring matter, in the case of the seal furs, was applied only to the fur portion, while the lambskins appear to have been dip-dyed. By dip-dyeing more

[1] J. R. Kanagy, J. Am. Leather Chem. Assoc. **33**, 352 (1938).

copper would be distributed on the leather than, would be if the dye bath was applied only to the fur surface.

Another reason the seal fur did not deteriorate as much as might be expected from the copper content, as compared to the Karakuls, may be that the oil tannage is inherently more resistant than the predominantly alum-type dressing of the lambskins.

The rate with which the metal salts are leached or removed from the lambskin pelts holds no relationship to the extent of deterioration. In other words the degree of the fixation of metal salts does not follow the aging. The results of water extraction of the metal salts from the skins are shown in table 7.

TABLE 7. *Aluminum, iron, and copper content of water-soluble extractives*

Constituent	Year of crop					
	Group I				Group II	
	1930	1931	1932	1933	1945	1946
	Percent	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*
Al_2O_3	0.387	0.129	0.290	0.112	0.295	0.165
Fe_2O_3	.006	.007	.046	.020	.080	.030
CuO	.065	.006	.030	.012	.058	.010
Total	0.458	0.142	0.366	0.144	0.433	0.205

VI. Summary

Aged crossbred Karakul lamb furs appear to have different physical properties than aged purebred Karakul lamb furs.

Purebred Karakul lamb furs in storage 15 years or more are considerably weaker in strength than similar pelts in storage approximately 4 years. Differences in strength between the age groups are not attributed to a change in tannage procedure over the consecutive years of processing.

Comparable losses in strength occurred in Alaskan seal furs after a storage period of 10 years.

Evidences of acid deterioration were not found.

The presence of small quantities of copper salts, which are introduced into the furs during dyeing, relates significantly with the extents of deterioration, and provides evidence that the metal acting catalytically promotes deterioration. Iron salts apparently do not promote deterioration appreciably.

The findings are in accord with a previously published report on the influence of copper and iron on vegetable-tanned leather deteriorated under accelerated conditions. The findings are novel in that the leathers described herein were of an alum or oil tannage, which were aged under normal shelf storage conditions.

The extents of deterioration are conveniently measured by determining either burst, breaking strength, or stitch tear (furrier-sewn) values. Losses in strength are accompanied by lower petroleum ether extractives, increased water solubles, increased soluble proteinaceous and other water-extractable organic matter.

WASHINGTON, June 19, 1950.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2142

Second Dissociation Constant of Succinic Acid from 0° to 50° C

By Gladys D. Pinching and Roger G. Bates

The second dissociation constant of succinic acid at intervals of 5 degrees from 0° to 50° C was determined from measurements of the electromotive force between hydrogen and silver-silver chloride electrodes in cells without liquid junction. The 19 solutions studied were divided into two series. The first series was made up of aqueous mixtures of sodium acid succinate, sodium succinate, and sodium chloride in the molar proportions 1:3:3. The solutions of the second series contained only sodium acid succinate and sodium chloride, in the molar ratio 1:1. The measurements of the second series were used to compute an approximate value for the first dissociation constant of the acid to be employed in correcting the results of the first series for the amount of free succinic acid formed by hydrolysis.

The second dissociation constant, K_2, for the temperature range studied can be expressed by

$$-\log K_2 = 1679.13/T - 5.7043 + 0.019153\ T,$$

where T is the temperature on the Kelvin scale. Thermodynamic quantities for the dissociation of acid succinate ion were calculated from the temperature coefficient of the second dissociation constant. At 25° C, the standard free-energy change was found to be 32,182 abs j mole^{-1}, and the changes of heat content and entropy for the dissociation process in the standard state were -450 abs j mole^{-1} and -109.4 abs j deg^{-1} mole^{-1}, respectively.

I. Introduction

The simultaneous dissociation of two or more acid groups of nearly equal strength complicates the determination of the dissociation constants of many dibasic and polybasic organic acids. Accurate values may indeed be impossible to obtain when the degree of dissociation is large in each overlapping step. Consequently, the thermodynamic quantities, such as heat-content and entropy changes, associated

with the ionization of acids with more than one carboxyl group have received little careful study.

The three dissociation constants, K_1, K_2, and K_3, of citric acid have been determined from electromotive-force measurements with hydrogen and silver-silver chloride electrodes in cells without liquid junction [1].[1] There was a considerable amount of overlapping, inasmuch as K_1/K_2 and K_2/K_3 are both about 44 at 25° C. It was possible, however, to facilitate the separation by choice of favorable buffer ratios and by a supplementary determination of the products K_1K_2 and K_2K_3 according to a method recently described [2].

These methods have now been applied to a determination of the second dissociation constant of succinic acid from 0° to 50° C at intervals of 5 degrees. The ratio K_1/K_2 for succinic acid, a dibasic acid, is about 26. Hence, the overlapping of these two dissociation steps is somewhat greater than that of the corresponding steps in the ionization of citric acid. The changes of free energy, heat content, entropy, and heat capacity accompanying the dissociation of one mole of acid succinate ion in the standard state of 0° to 50° C have been evaluated from the dissociation constant and its temperature coefficient.

II. Experimental Procedures and Results

Sodium acid succinate and sodium succinate were both prepared from reagent-grade succinic acid and standardized sodium hydroxide solution. The sodium succinate was recrystallized twice from water and dried at 125° C. Ignition of the purified salt to sodium carbonate indicated a purity of 99.99 ± 0.02 percent. The sodium acid succinate was analyzed by titration with standard alkali and found to contain a small amount of free acid. It was therefore recrystallized from water with the addition of the amount of sodium hydroxide required to bring the salt to the proper composition. The anhydrous salt was obtained by drying at 110° C. The purity was found by titration to be 100.03 ± 0.02 percent. Sodium chloride, practically free from bromide, was recrystallized from water. The product contained less than 0.001 percent of bromide, as shown by the uranine test [3].

Neither of the succinate salts is appreciably hygroscopic. Exposed at 25° C to air of 45-percent relative humidity, sodium succinate took up only 0.012 percent of moisture in 10 days, and the acid salt gained only 0.006 percent in 24 hours.

Four stock solutions were prepared by weight from the purified salts and water having a specific conductance of approximately 0.5×10^{-6} reciprocal ohm. Each stock solution was diluted with water to form a series of four or five cell solutions. Dissolved air was removed by passage of nitrogen through each solution before the cells were filled.

The hydrogen electrodes were formed by depositing platinum black upon bases of platinum foil. Each foil was electrolyzed for 2 minutes at a current of 300 ma in a 3-percent solution of chloroplatinic acid that was 0.002 *M* with respect to lead acetate. The preparation of the silver chloride electrodes has been described [4]. The cells were those heretofore employed for studies of systems of weak acids and their salts in which the solubility of silver chloride is negligible [5]. Each cell, with electrodes in place, was twice filled with solution and emptied by pressure of pure hydrogen before the final portion was admitted. The cells were allowed about 5 hours to attain initial equilibrium at 25° C, after which the temperature was lowered overnight to near 0° C. Measurements of the electromotive force were made from 0° to 25° C on the second day and from 25° to 50° C on the third day. At the conclusion of the series, final values at 25° were obtained. These usually agreed with the initial values within 0.00010 v.

The platinum-hydrogen electrodes were mounted just beneath the surface of the solution, and the mean total pressure at each electrode was assumed to equal that of the atmosphere. Hills and Ives [6] have recently found that the effective pressure, that is, the pressure that fixes the concentration of dissolved hydrogen in equilibrium with the electrode, is determined by the depth of the jet as well as by the gas-phase pressure. The solution becomes saturated with gas at the jet level and is thoroughly mixed by the bubbling. The free surface apparently is slow to relieve supersaturation. The excess pressure is, according to Hills and Ives, about two-fifths of the hydrostatic pressure at the jet. The jet in the cells used here is located about 5 cm below the surface of the solution. At a hydrogen pressure of 1 atm, a change of 1 mm Hg (13.6 mm water) in the partial pressure of hydrogen changes the potential of the hydrogen electrode by about 16 μv at 25° C and 21 μv at 50° C. Hence, a jet depth of 5 cm, or an excess pressure of 2 cm water, would cause an error of 0.00002 v at 25° C and 0.00003 v at 50° C. This error is somewhat less than the observed differences between duplicate cells. No attempt has been made to apply a correction for this effect. It is presumably largely compensated by a comparable error in the values for the standard potential of the hydrogen-silver chloride cell, which were derived from measurements of hydrochloric-acid solutions contained in cell vessels of a similar design [7].

The electromotive force of the cell

Pt; $H_2(g, 1\text{ atm})$, NaHSuc (m), $Na_2Suc(3m)$, NaCl $(3m)$, AgCl(s); Ag

from 0° to 50° C is listed in table 1 (series 1). Each recorded value is the mean of the potentials between two pairs of electrodes in the same cell. Here, as elsewhere in this paper, $Suc^=$ is written for succinate anion ($^-OOCCH_2CH_2COO^-$) and $HSuc^-$ for acid succinate ion. Similarly, succinic acid will be abbreviated H_2Suc. Indication of the charges of hydrogen and chloride ions will be omitted. The data of table 1 (series 2) are for the cell

Pt; $H_2(g, 1\text{ atm})$, NaHSuc (m), NaCl (m), AgCl (s); Ag.

[1] Figures in brackets indicate the literature references at the end of this paper.

TABLE 1. *Electromotive force of hydrogen-silver chloride cells containing succinate-chloride solutions at 0° to 50° C.*

m_{NaHSuc}	Temperature, °C										
	0	5	10	15	20	25	30	35	40	45	50
Series 1. NaHSuc:Na_2Suc:NaCl=1:3:3											
0.05012	0.60112	0.60417	0.60722	0.61036	0.61341	0.61645	0.61977	0.62292	0.62612	0.62939	0.63278
.04951	.60142	.60447	.60750	.61066	.61376	.61686	.62002	.62307	.62638	.62958	.....
.04555	.60364	.60674	.60988	.61303	.61622	.61935	.62259	.62583	.62913	.63243	.63574
.03985	.60696	.61012	.61335	.61648	.61977	.62304	.62635	.62972	.63313	.63647	.63982
.03715	.60891	.61214	.61536	.61869	.62195	.62526	.62840	.63191	.63535	.63874	.64233
.03556	.60998	.61325	.61655	.61988	.62323	.62649	.62992	.63331	.63677	.64023	.64371
.02511	.61896	.62242	.62594	.62930	.63306	.63648	.64020	.64380	.64745	.65112	.65477
.02509	.61900	.62248	.62601	.62956	.63306	.63666	.64020	.64384	.64752	.65121	.65483
.02493	.61925	.62278	.62625	.62981	.63329	.63675	.64054	.64415	.64778	.65143	.65530
.014886	.63277	.63655	.64035	.64421	.64806	.65189	.65579	.65973	.66369	.66767	.67165
.014853	.63291	.63671	.64050	.64433	.64807	.65195	.65599	.65988	.66381	.66777	.67199
.010184	.64262	.64662	.65065	.65467	.65875	.66267	.66695	.67105	.67520	.67925	.68330
.007573	.65035	.65456	.65878	.66300	.66725	.67139	.67578	.68003	.68441	.68871	.69263
.005074	.66083	.66523	.66965	.67408	.67849	.68287	.68740	.69196	.69652	.70113	.70574
Series 2. NaHSuc:NaCl=1:1											
0.10000	0.55454	0.55673	0.55898	0.56121	0.56350	0.56557	0.56781	0.57007	0.57225	0.57445	0.57663
.07296	.56248	.56484	.56724	.56957	.57194	.57434	.57666	.57899	.58136	.58377	.58610
.05012	.57187	.57443	.57702	.57956	.58223	.58469	.58727	.58986	.59244	.59492	.59734
.03155	.58345	.58627	.58909	.59188	.59484	.59730	.60036	.60323	.60599	.60880	.61154
.019390	.59569	.59874	.60183	.60484	.60793	.61069	.61390	.61695	.61993	.62294	.62590

III. Calculation of the Dissociation Constant

When the two acid groups of a dibasic acid are of widely different strengths, the second dissociation constant can readily be computed from a single series of measurements of buffer solutions such as those of series 1.[2] To accomplish this, the mass-action expression is combined with the equation that relates the electromotive force to the activities of the reactants and products of the cell reaction. When exact values of the activity coefficients are unknown, the semi-empirical Hückel formula [14] can be used to facilitate the determination of the true thermodynamic dissociation constant, K, by extrapolation to zero ionic strength. The resulting expression for the apparent dissociation constant, K', (obtained with the use of inexact activity coefficients) is

$$-\log K_2' \equiv -\log K_2 - \beta\mu = \mathrm{pwH} - \log \frac{m_{Suc^=}}{m_{HSuc^-}} + \frac{2A\sqrt{\mu}}{1+Ba^*\sqrt{\mu}}, \quad (1)$$

where A and B are constants of the Hückel equation,[3] a^* and β are adjustable parameters, and μ is the ionic strength.

The quantity pwH is defined by

$$\mathrm{pwH} \equiv -\log (f_H f_{Cl} m_H) = (E - E^\circ)\boldsymbol{F}/(2.3026RT) + \log m_{Cl}, \quad (2)$$

where f is an activity coefficient on the molal scale, E and E° are the electromotive force and standard potential,[3] respectively, of the hydrogen-silver chloride cell, $\boldsymbol{F}$ is the faraday, R the gas constant, and T the temperature on the Kelvin scale.

Equation 1 also applies when the two dissociation steps overlap, as do those of succinic acid. However, the terms in m represent the actual ionic molalities and cannot be assumed to be equal to the stoichiometrical molalities of the primary and secondary succinate salts. Instead, the detailed ionic composition of each solution must be considered.

1. The Hydrogen-Ion Correction

In order to compute the actual concentrations of the succinate ions and the second term on the right of eq 1, the equilibria among these ions must be considered. The equilibria involving acid succinate ion in aqueous solution may be represented by

$$HSuc^- = H^+ + Suc^= \quad (3)$$

and

$$2\ HSuc^- = H_2Suc + Suc^=. \quad (4)$$

In series 1, where the stoichiometric molality of sodium succinate is $3m$ and that of acid succinate m,

$$m_{HSuc^-} = m - m_H - 2\ m_{H_2Suc} \quad (5)$$

and

$$m_{Suc^=} = 3\ m + m_H + m_{H_2Suc}. \quad (6)$$

The ionic strength of each solution of series 1 is given by

$$\mu = 13\ m + 2\ m_H + m_{H_2Suc}. \quad (7)$$

[2] For the application of this method to the determination of second dissociation constants when overlapping is negligible ($K_1/K_2 > 1{,}000$), and for the derivation of eq 1, the reader is referred to the following papers on the respective acids: carbonic [8, 9], malonic [10], oxalic [11, 12], phenolsulfonic [4], phosphoric [5, 13].

[3] The constants A and B at 0° to 100° C are given in [15] and E° and $2.3026RT/\boldsymbol{F}$ in abs v in [16].

The hydrogen-ion concentration is small relative to m and is estimated with sufficient accuracy by

$$-\log m_H = \text{pwH} + \log f_H f_{Cl} = \text{pwH} - \frac{2A\sqrt{\mu}}{1+Ba^*\sqrt{\mu}}. \quad (8)$$

The quantity $f_H f_{Cl}$ represents the square of the mean activity coefficient of hydrochloric acid in the succinate buffer solutions. Although this is unknown, $f_H f_{Cl}$ in pure hydrochloric acid solutions at ionic strengths of 0.05 to 0.2 [7, 17] is represented by the last term of eq 8 with a^* values of 5.8 to 6.2. Also $f_H f_{Cl}$ in pure solutions of sodium chloride of the same concentrations [18] can be expressed similarly with values of 5 to 5.6 for a^*. Although the correct value of a^* thus appears to lie near 6 in solutions containing chiefly sodium salts, a change from 4 to 7 produces a negligible change in the second dissociation constant, for the hydrogen-ion correction is small. Inasmuch as the best extrapolation was obtained when a value of 7 was used for a^* in eq 1, it was found convenient to employ this value in eq 8 as well.

2. Correction for the First Dissociation Step

As a result of overlapping of the first and second acid groups, a part of the acid succinate ion is converted into molecular succinic acid. The concentration of succinic acid in the solutions of series 1 was only about 2 percent of m, since the excess of the normal succinate over the acid salt was effective in suppressing reaction 4. Because the residual overlapping was small, a precise knowledge of the first dissociation constant, K_1, and the activity coefficients was not needed to obtain sufficiently accurate figures for the molality of succinic acid in eq 5 and 6. If $f_H f_{HSuc^-}$ is approximately equal to $f_H f_{Cl}$ in the same solutions, a not unreasonable assumption, we have from the definition of pwH (eq 2) and the mass law for the first dissociation step,

$$m_{H_2Suc} = m_{HSuc^-}/(K_1 \text{ antilog pwH}). \quad (9)$$

For an accuracy of ± 0.002 in log K_2, it was necessary to know log K_1 only to ± 0.05. Values in the literature are quite adequate at and near 25° C but do not extend to 0° and 50° C. Hence, five solutions of sodium acid succinate and sodium chloride (series 2) were studied in order to obtain an approximate value of K_1 throughout the temperature range. For the solutions of series 2,

$$-1/2 \log (K_1K_2)' = \text{pwH} - 1/2 \log \frac{mx + m_H}{mx - m_H} + \frac{A\sqrt{\mu}}{1+Ba^*\sqrt{\mu}}, \quad (10)$$

where the prime indicates again the "apparent" value and

$$x \equiv \frac{\sqrt{4k_2/k_1}}{1+\sqrt{4k_2/k_1}}. \quad (11)$$

The symbols k_1 and k_2 represent the concentration constants. The details of the application of eq 10 and 11 to the calculation of K_1K_2 have been given elsewhere [1, 2] and need not be repeated. The solution of eq 1 and 10 was most conveniently accomplished by successive approximations, of which only two were usually necessary. The extrapolation of $-1/2 \log (K_1K_2)'$ is shown in figure 1. The limiting values of $-1/2 \log K_1K_2$ and of $-\log K_1$ at each temperature are given in table 2. Inasmuch as these are based upon too few points to provide an unambiguous extrapolation, they are to be regarded as approximate. Nevertheless, log K_1 is probably correct to ± 0.01 and, hence, more than adequate for the purpose.

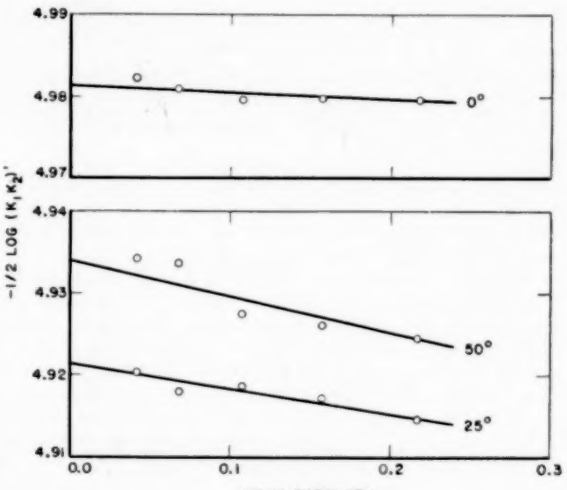

FIGURE 1. *Determination of the product K_1K_2 by extrapolation to zero ionic strength.*

TABLE 2. *Values of log K_1K_2 and log K_1 at 0° to 50° C*

Temperature	$-1/2 \log K_1K_2$	$-\log K_1$	Temperature	$-1/2 \log K_1K_2$	$-\log K_1$
°C			°C		
0	4.981	4.28$_8$	30	4.921	4.20$_5$
5	4.963	4.26$_6$	35	4.921	4.19$_4$
10	4.948	4.24$_7$	40	4.924	4.19$_4$
15	4.937	4.23$_2$	45	4.930	4.19$_5$
20	4.931	4.22$_5$	50	4.934	4.18$_8$
25	4.921	4.20$_5$			

3. The Extrapolation

Figure 2 demonstrates the effect of three different values of a^* (eq 1) upon pK_2' (that is, $-\log K_2'$) at 25° C. From top to bottom of the figure, the three solid lines correspond to a^* values of 6, 7, and 8, respectively. The upper dashed line, computed for $a^*=7$, illustrates the effect of omitting entirely the correction for the first dissociation, that is, of ignoring m_{H_2Suc} in eq 5 and 6. Although K_1/K_2 is only 26, it appears that the 3:1 buffer ratio has materially reduced the extent of overlapping, for the K_1 correc-

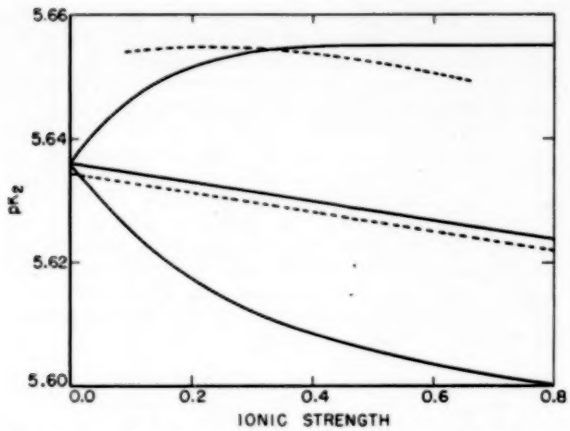

FIGURE 2. *Effect of a^* on the course of the extrapolation line.*

Solid lines represent (top to bottom) $a^*=6$, 7, and 8. The dashed lines show the course of the extrapolation when the correction for the overlapping first dissociation is ignored (upper dashed line) or made with a value of log K_1 that is in error by 0.05 unit (lower dashed line).

tion contributes only about 0.02 to log K_2'. The lower dashed line indicates the change of position of the center extrapolation line that would result from an error of 0.05 in log K_1.

When a^* was taken to be 7, log K_2' was found to be a linear function of ionic strength at each temperature, as shown in figure 3, although the slopes of the lines changed considerably with temperature. These slopes ($-\beta$ in eq 1) were measured and log K_2 calculated from each experimental point with the aid of eq 1. The average values are given in the second column of table 3, together with the mean deviation from the average at each temperature. The corresponding K_2 appears in the last column. The uncertainty in log K_2 is believed to be less than ± 0.003 unit.

TABLE 3. *Values of log K_2 and K_2 from 0° to 50° C*

Temperature	$-\log K_2$	$K_2\times10^3$	Temperature	$-\log K_2$	$K_2\times10^3$
°C			°C		
0	$5.674_1\pm0.0013$	2.12	30	$5.641_6\pm0.0014$	2.28
5	$5.660_8\pm .0012$	2.18	35	$5.647_7\pm .0011$	2.25
10	$5.649_8\pm .0009$	2.24	40	$5.654_0\pm .0012$	2.22
15	$5.642_5\pm .0011$	2.28	45	$5.669_3\pm .0013$	2.14
20	$5.639_0\pm .0008$	2.30	50	$5.680_3\pm .0026$	2.09
25	$5.635_4\pm .0012$	2.31			

Table 4 is a summary of the published values for the second dissociation constant grouped according to the method used for the determination. The constant found by Jones and Soper [33] should be most nearly comparable with that reported here, for their measurements were also made with a hydrogen-silver chloride cell without liquid junction. However, a different method of calculation was used. Unfortunately, no electromotive-force data are listed in their paper, so it is impossible to explain the differences of about 0.03 and 0.04 between their log K_2 and that of this investigation at 25° and 50° C, respectively.

TABLE 4. *Summary of determinations of log K_2*

Reference	Method *	Temperature	$-\log K_2$
	CONDUCTANCE		
		°C	
1908, Chandler [19]		25	5.57
1921, Duboux [20]	(Recalculation)	25	5.69
	DISTRIBUTION		
1908, McCoy [21]		25	5.37
1908, Chandler [19]		25	5.66
	REACTION RATE		
1927, Duboux and Frommelt [34]	Decomposition of diazoacetic ester.	25	5.55
	SOLUBILITY		
1915, Datta and Dhar [35]	Of CO_2	25	5.54
1926, Larsson [36]	Of benzoic acid	25	5.6
	ELECTROMOTIVE FORCE, CELLS WITH LIQUID JUNCTION		
1922, Larsson [22]	Q-cal	18	5.620
1924, Auerbach and Smolczyk [23]	Q-cal	20	5.401
1925, Britton [24]	H-cal	18	5.28
1926, Mizutani [25]	H-cal	18	5.43
1928, Simms [26]	H-cal	25	5.570
1928, Kolthoff and Bosch [27]	H-cal	18	5.57
1928, Gane and Ingold [28]	H-cal	25	5.347
1929, Ölander [29]	H-Q; Q at 18° C	50	5.730
1931, Gane and Ingold [30]	H-cal	25	5.478
1934, Ashton and Partington [31]	H-cal	25	5.322
1936, German and Vogel [32]	Q-cal	25	5.590
	ELECTROMOTIVE FORCE, CELLS WITHOUT LIQUID JUNCTION		
1936, Jones and Soper [33]	H-AgCl	25	5.607
		50	5.688
1950, This investigation	H-AgCl	20	5.729
		25	5.639
		50	5.686

* H = hydrogen electrode, Q = quinhydrone, cal = calomel.

IV. Thermodynamic Functions

The dissociation constants given in table 3 are adequately represented by either

$$-\log K_2 = 1679.13/T - 5.7043 + 0.019153\ T \quad (12)$$

or

$$-\log K_2 = 3756.1/T - 79.2716 + 29.223 \log T. \quad (13)$$

The mean deviation of the observed log K_2 from the calculated value is 0.0011 for eq 12 and 0.0018 for eq 13. The first is the form suggested by Harned and Robinson [37]; the second has been used especially by Everett and Wynne-Jones [38]. If eq 12 and 13 were used solely for the interpolation of K_2 at various temperatures, either would serve the purpose. Likewise, ΔF°, the standard free energy of dissociation, is readily obtained from the dissociation constant at each temperature. However, some expression of the variation of log K_2 with temperature is necessary in order to calculate, over the entire range, the quantities ΔH°, ΔS°, and ΔC_p°, which are

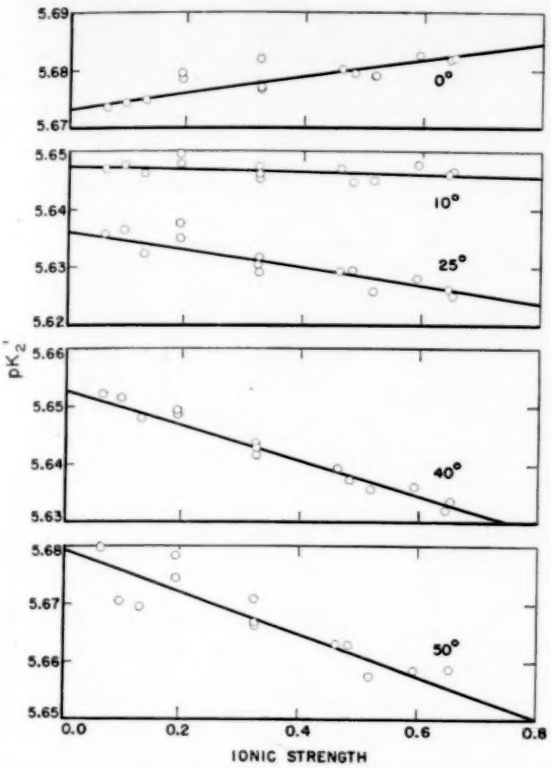

FIGURE 3. *Extrapolation lines from which $-\log K_2$ (pK_2) was obtained at 0°, 10°, 25°, 40°, and 50° C.*

respectively the changes of heat content, entropy, and heat capacity for the dissociation of 1 mole of acid succinate ion in the standard state. By differentiation of eq 12 and 13 with respect to temperature, the following expressions for $d \log K_2/dT$ are obtained:

$$d \log K_2/dT = 1679.13/T^2 - 0.019153 \qquad (12a)$$

and

$$d \log K_2/dT = 3756.1/T^2 - 12.6913/T. \qquad (13a)$$

Since the two equations yield somewhat different temperature coefficients, they will also furnish different values for these three thermodynamic functions. The difference is not large near the middle of the temperature range but is appreciable at the ends, as can be seen from the following comparison of ΔH° computed from the two equations:

Temperature	Eq 12	Eq 13
°C	*j mole*$^{-1}$	*j mole*$^{-1}$
0	4, 787	5, 541
25	−450	−533
50	−6, 146	−6, 607

Equation 12 rests upon the observation that the electromotive force of cells of the type used here (and hence ΔF°) is closely a quadratic function of temperature. It implies that ΔC_p° is a linear function of temperature. Equation 13 is based upon a presumed linear variation of ΔH° with T, and hence a value of ΔC_p° that is constant over the entire temperature range. For this reason, eq 12 is sometimes considered theoretically less acceptable than eq 13. Nevertheless, the authors are of the opinion that the fit of the observed data provides the most valid justification for choice of one empirical or semi-empirical equation over another. Equation 12 is usually found to represent the observed dissociation constants for weak acids and bases more closely than eq 13. Hence, the former was used to compute the thermodynamic functions given in table 5 by application of the usual formulas.

TABLE 5. *Thermodynamic quantities for the dissociation of acid succinate ion*

Temperature	ΔF°	ΔH°	ΔS°	ΔC_p°
°C	*abs j mole*$^{-1}$	*abs j mole*$^{-1}$	*abs j deg*$^{-1}$ *mole*$^{-1}$	*abs j deg*$^{-1}$ *mole*$^{-1}$
0	29, 675	4, 787	−91. 1	−200
5	30, 140	3, 776	−94. 8	−204
10	30, 623	2, 747	−98. 4	−208
15	31, 125	1, 690	−102. 1	−211
20	31, 645	633	−105. 8	−215
25	32, 182	−450	−109. 4	−219
30	32, 739	−1, 553	−113. 1	−222
35	33, 314	−2, 674	−116. 8	−226
40	33, 907	−3, 813	−120. 4	−230
45	34, 519	−4, 971	−124. 1	−233
50	35, 148	−6, 146	−127. 8	−237

The estimated uncertainty of 0.003 in $\log K_2$ implies a corresponding uncertainty of about 18 j mole^{-1} in ΔF°. If the values of $\log K_2$ at 0° and 50° C were each in error by 0.003 unit and the errors were of opposite sign, ΔH° at 25° would be altered by about 300 j mole^{-1} and ΔS° by about 1 j deg^{-1} mole^{-1}. Similarly ΔC_p° at 25° C can be assigned an uncertainty of about 12 j deg^{-1} mole^{-1}.

Cottrell and Wolfenden [39] have reported a calorimetric determination of the heats of dissociation of the two steps of succinic acid from 5° to 30° C. By combination of their data with ΔF° taken from the work of Jones and Soper [33], they computed the entropy of dissociation, ΔS°. The change of heat capacity, ΔC_p°, was obtained from the temperature variation of ΔH°. The value of Cottrell and Wolfenden for the entropy change has been adjusted to correspond with the ΔF° of this investigation, which seems more certain than the earlier result. The values at 25° C are:

	Cottrell and Wolfenden	Pinching and Bates
ΔH°	176 j mole^{-1}	−450 j mole^{-1}
ΔS°	−107.4 j deg^{-1} mole^{-1}	−109.4 j deg^{-1} mole^{-1}
ΔC_p°	−218 j deg^{-1} mole^{-1}	−219 j deg^{-1} mole^{-1}

904428—50——6

The two values for the heat of dissociation differ by the rather large amount of 626 j. Cottrell and Wolfenden state that their result is subject to considerable uncertainty, for it represents a difference between two experimental quantities and also rests upon a rather arbitrary extrapolation. Differences of extrapolation often do not affect the temperature coefficient, and it is perhaps significant that practically identical values of $\triangle C_p^\circ$, that is $d(\triangle H^\circ)/dT$, were obtained by the two methods. Cottrell and Wolfenden found $\triangle H^\circ$ to pass through zero at 26° C, in reasonably good agreement with 23° C at which the value of $-\log K_2$ given in table 3 reaches a minimum.

V. References

[1] R. G. Bates and G. D. Pinching, J. Am. Chem. Soc. **71,** 1274 (1949).
[2] R. G. Bates, J. Am. Chem. Soc. **70,** 1579 (1948).
[3] G. D. Pinching and R. G. Bates, J. Research NBS **37,** 311 (1946) RP1749.
[4] R. G. Bates, G. L. Siegel, and S. F. Acree, J. Research NBS **31,** 205 (1943) RP1559.
[5] R. G. Bates and S. F. Acree, J. Research NBS **30,** 129 (1943) RP1524.
[6] G. J. Hills and D. J. G. Ives, Nature **163,** 997 (1949).
[7] H. S. Harned and R. W. Ehlers, J. Am. Chem. Soc. **55,** 2179 (1933).
[8] D. A. MacInnes and D. Belcher, J. Am. Chem. Soc. **55,** 2630 (1933).
[9] H. S. Harned and S. R. Scholes, Jr., J. Am. Chem. Soc. **63,** 1706 (1941).
[10] W. J. Hamer, J. O. Burton, and S. F. Acree, J. Research NBS **24,** 269 (1940) RP1284.
[11] H. S. Harned and L. D. Fallon, J. Am. Chem. Soc. **61,** 3111 (1939).
[12] G. D. Pinching and R. G. Bates, J. Research NBS **40,** 405 (1948) RP1885.
[13] L. F. Nims, J. Am. Chem. Soc. **55,** 1946 (1933).
[14] E. Hückel, Physik. Z. **26,** 93 (1925).
[15] G. G. Manov, R. G. Bates, W. J. Hamer, and S. F. Acree, J. Am. Chem. Soc. **65,** 1765 (1943).
[16] R. G. Bates and G. D. Pinching, J. Research NBS **42,** 419 (1949) RP1982.
[17] H. S. Harned and B. B. Owen, The Physical chemistry of electrolytic solutions, chapter 11 (Reinhold Publishing Corp., New York, N. Y., 1943).
[18] H. S. Harned, J. Am. Chem. Soc. **57,** 1865 (1935).
[19] E. E. Chandler, J. Am. Chem. Soc. **30,** 694 (1908).
[20] M. Duboux, J. chim. phys. **19,** 179 (1921).
[21] H. N. McCoy, J. Am. Chem. Soc. **30,** 688 (1908).
[22] E. Larsson, Z. anorg. Chem. **125,** 281 (1922).
[23] F. Auerbach and E. Smolczyk, Z. physik. Chem. **110,** 65 (1924).
[24] H. T. S. Britton, J. Chem. Soc. **127,** 1896 (1925).
[25] M. Mizutani, Z. physik. Chem. **118,** 318 (1926).
[26] H. S. Simms, J. Phys. Chem. **32,** 1121 (1928).
[27] I. M. Kolthoff and W. Bosch, Rec. trav. chim. **47,** 861 (1928).
[28] R. Gane and C. K. Ingold, J. Chem. Soc., 1594 (1928).
[29] A. Olander, Z. physik. Chem. **A144,** 49 (1929).
[30] R. Gane and C. K. Ingold, J. Chem. Soc., 2153 (1931).
[31] H. W. Ashton and J. R. Partington, Trans. Faraday Soc. **30,** 598 (1934).
[32] W. L. German and A. I. Vogel, J. Am. Chem. Soc. **58,** 1546 (1936).
[33] I. Jones and F. G. Soper, J. Chem. Soc., 133 (1936).
[34] M. Duboux and J. Frommelt, J. chim. phys. **24,** 245 (1927).
[35] A. K. Datta and N. Dhar, J. Chem. Soc. **107,** 824 (1915).
[36] E. Larsson, Z. anorg. Chem. **155,** 247 (1926).
[37] H. S. Harned and R. A. Robinson, Trans. Faraday Soc. **36,** 973 (1940).
[38] D. H. Everett and W. F. K. Wynne-Jones, Trans. Faraday Soc. **35,** 1380 (1939).
[39] T. L. Cottrell and J. H. Wolfenden, J. Chem. Soc., 1019 (1948).

WASHINGTON, May 4, 1950.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2143

Methods of Sieve Analysis With Particular Reference to Bone Char[1]

By Frank G. Carpenter[2] and Victor R. Deitz

The procedure for separating particle sizes of solid adsorbents by sieving has been studied in detail because of the influence of the particle size on adsorbent properties. The investigation was conducted chiefly with sieve openings in the range between U. S. Standard Sieves No. 8 and No. 80. The largest source of error is in the testing sieves themselves. This is due to the tolerances permitted by the present specifications. It is feasible to calibrate testing sieves by the use of a calibrated sample of spherical glass beads and thus obtain the opening that is effective in sieving. The calibration of testing sieves in this manner can lead to reproducible sieve analyses by different laboratories. A simple procedure is proposed to determine the uniformity of sieve openings and, thereby, to furnish a criterion for the discard of distorted sieves. An analysis with seven Ro-Tap machines indicated that, in general, best results are obtained when the Ro-Tap is operated at 115 taps/min of the knocker mechanism. The other variables concerned with shaking that were examined are of minor importance and need not be rigidly controlled. As first choice, the weight of the sample should be between 100 and 150 g. The shaking time should be adjusted to the weight and the particle size distribution of the sample according to relationships developed.

I. Introduction

The adsorbent properties of bone char and other materials are greatly influenced by particle size. Sieve analysis is one of the basic tests for measuring the particle size of all powdered and granular materials. In sugar refining and many other industries great importance is ascribed to sieve analyses in spite of the fact that their reproducibility may be extremely poor, in comparison with other measurements such as volume, mass, or chemical composition. Differences in the results of sieve analyses of

[1] This investigation was sponsored as a joint research project undertaken by the United States Cane Sugar Refiners and Bone Char Manufacturers, a greater part of the refining industry of the British Commonwealth, Belgium and the National Bureau of Standards.

[2] Research Associate at the National Bureau of Standards, representing the cooperating manufacturers.

the order of magnitude of 10 percent are frequently reported between different laboratories, while differences of 1 or 2 percent are not uncommon with the same operator and the same sieving equipment. In view of the importance of sieve analysis and of the poor reproducibility when compared to other measurements, an investigation was made of the sources of errors and of the steps that might be taken to reduce the errors to a minimum. Previous work on this subject is not considered adequate [1, 2].[3]

The sources of errors in a sieve analysis can be classified under three general headings: The sieves, the method of shaking, and the sample. The sieves considered are the standard 8-in. laboratory testing sieves. The nominal sizes of the openings of the standard sieves ($\sqrt{2}$ series) are in a geometric series with a fixed ratio of the square root of 2. For closer sizing, an additional sieve is provided between each pair in the standard series thus forming the closest-sizing series ($\sqrt[4]{2}$ series). The methods of shaking the sieves that were studied were by hand and by the use of the Ro-Tap [4] machine. Although there are other shaking devices, the Ro-Tap machine, which is one of those commonly employed, was the only type studied. The majority of the samples studied were bone chars from various cane sugar refineries throughout the world. Other materials included were crystalline refined sugar, glass, sand, iron filings, and granular bismuth. The particle shapes included were spherical, irregular, and long needle-like granules. The particle sizes ranged from those passing a No. 4 sieve [5] to those retained on a No. 270 sieve. These correspond to particle diameters of 0.476 to 0.0053 cm, or 0.187 to 0.0021 in. Most of the work was conducted in the sieve size range No. 8 to No. 80.

The application of the results of a sieve analysis to the interpretation of physical properties presents many difficult problems, and these are not considered here in all their aspects. DalleValle [3] has considered many phases of sieve analysis, but unfortunately these were not all examined critically by him. Hatch [4] considered the relations between weight-size and number-size distribution and also the various methods of evaluating the average particle size. The relation between the size of the sieve openings and the average diameter of irregularly shaped particles that will pass through them has not been adequately investigated. The method of attack emphasized in this investigation makes use of glass spheres to define the effective openings of sieves.

1. Method of Reporting Sieve Analyses

Sieve analyses are usually reported as percentage by weight of the total sample that passes a certain sieve and that is retained on another, for example, "percentage passing No. 20 and retained on No. 30". This is frequently shortened for convenience to, for example, "20 to 30 fraction" or "20 to 30 mesh". The use of the term *mesh* in this sense is to be avoided [5]. For some comparisons, sieve analyses are reported as cumulative percentage finer or coarser than a certain sieve.

When evaluating differences between sieve analyses, the differences are expressed in percentage of the total sample for a certain sieve fraction. If the difference (or deviation) is compared to the mean value, then a percentage of the percentage of the total sample is obtained. Such a terminology is awkward, and there is used instead the phrase "coefficient of variation", which is expressed in percentage.

In evaluating the variation among several sieve analyses, the standard deviation is computed for each sieve in the usual manner. The standard deviations for the different sieve fractions are not strictly comparable for reasons that will be apparent, although they have been averaged in some cases to give a single figure as a measure of the reproducibility of the sieve analyses.

II. Sieves

One of the largest sources of variation in sieve analyses is in the testing sieves themselves. The wire size, average opening, and uniformity of opening are specified [6] with a small variation allowed in each for manufacturing tolerances. The wire size is of minor consequence, because it does not directly influence the size of particle that will be retained on the sieve.

1. Variation in Average Opening

The variation in average opening allowed by the specifications may seem stringent from a manufacturing point of view; but it has an important effect on sieve analyses, since the variation ranges from ± 3 percent for sieves No. 16 and coarser to ± 7 percent for sieves No. 200 and finer. For example, the variation allowed for sieves No. 20 and 30 ($\sqrt{2}$ series) is ± 5 percent of the nominal opening. The No. 20 sieve, whose nominal opening is 0.084 cm, thus may have average openings ranging between 0.080 and 0.088 cm. Similarly, No. 30 sieves of nominal opening 0.059 cm may have average openings somewhere between 0.056 and 0.062 cm. When these two sieves, both uniformly woven, are used to define a sieve fraction, the particle sizes that are retained between the two sieves depend on the openings in both sieves. The mean value of the openings can, therefore, indicate the probable spread in particle sizes. The maximum spread using these two sieves is 0.032 cm, and the minimum is 0.018 cm. It is noted that the maximum would be *twice* that of the minimum.

This type of error is much larger when sieves of the $\sqrt[4]{2}$ series are used. This is best explained on a percentage basis. The nominal opening of each sieve is only 19 percent larger than that of the next smaller. When sieves No. 200 or finer are used, and the average opening of the larger sieve is oversize to

[1] Figures in brackets indicate the literature references at the end of this paper.
[3] Manufactured by W. S. Tyler Co., Cleveland, Ohio.
[5] All sieve numbers, unless otherwise noted, refer to the U. S. Standard series.
[6] Identical specifications are issued by the American Society for Testing Materials (ASTM) [6], The American Standards Association (ASA) [7], and the National Bureau of Standards [5].

the limit of the specifications, and the average opening of the smaller sieve is undersize to the limit of the specifications, then this 19 percent becomes approximately 19+7+7=33 percent. In the converse case, the figure becomes 19−7−7=5 percent. If the size distribution by weight of the material being sieved is approximately constant, then the amount of material remaining between the two sieves in the first case would be nearly *seven* times as much by weight as in the second case, the sieves being uniformly woven. Sieves that did not conform to specifications might give variations many times greater than this. With careful selection sieves can be held to tolerances much closer than the specifications, with proportionally better results. For the best results it would be necessary to calibrate individually each sieve to find the actual effective openings of that particular sieve. When reproducibility alone is desired and accuracy of measurement of particle size of minor importance, sieves may be carefully matched.

2. Calibration of Sieves

Testing sieves are calibrated from measurements with a projection microscope (see [5]). For a small fee, the National Bureau of Standards calibrates sieves by this method. This service is generally available only for new sieves. In a calibration of this type, the thickness of the wires is measured, and the number of openings per inch is counted usually for a distance of at least 6 in. The average opening is then determined by dividing the difference between the total distance measured and the sum of all wire diameters contained in this distance by the number of openings.

It is to be noted that if the size of the openings is not quite uniform, then the larger openings are the effective ones, and the average opening is almost meaningless. Weber and Moran [8] suggested an absolute microscopic calibration of testing sieves consisting in measuring a large number of the openings and determining the effective size of the openings by use of an empirical relation between the statistical parameters and the effective opening. This calibration method may be too long and involved to be of practical value. However, when particles to be sieved are fairly symmetrical in shape, the effective opening of a sieve can be easily determined by measuring the size of spherical particles that will pass. This type of measurement is best done by means of calibrated samples of material of known particle-size distribution, determined by some means other than sieving. This method has the advantage of being quickly done without special equipment by anyone familiar with sieving procedures. It also has the advantage that used sieves may be checked periodically to determine whether wear or deformation has occurred to an extent sufficient to make them unsuitable.

The material of the calibrated sample should be hard enough to eliminate completely any question of abrasion. The best shape of the particles for calibration purposes is spherical, because the diameter of the particles can be measured by microscopic means, and there is no doubt about the orientation of the particle as it passed the sieve opening. Glass beads of the type used for highway markings are suitable for this use when properly selected.[7]

3. Glass Beads as Reference Samples for Calibration of Sieves

Spherical glass beads ranging in size from those passing a No. 20 sieve to those retained on a No. 100 sieve were used to explore the possibilities of calibrating testing sieves. With this sample, sieves from No. 25 through No. 80 can be calibrated. A sorting procedure was developed to remove all misshapen particles from the material as received. The material was then divided into smaller quantities by a sample divider of the riffle type. One of these samples was sieved into 10 fractions so beads of nearly the same size would be grouped together for ease of measurement. Four samples of about 100 beads were taken from each of these 10 fractions and a microscope slide prepared of each. Twenty-five beads from each slide were measured in air with a micrometer microscope.

From the measured diameters of 1,000 particles, the particle-size distribution was evaluated. At least 10,000 particles are desirable for accurate results. However, as the object of this experiment was only to explore the feasibility of such a calibration sample, this refinement was not considered necessary. Figure 1 is a plot of the particle-size distribution for the glass beads. The percentage by weight finer than any specific size is plotted against that size of bead as determined microscopically.

4. Procedure in Using the Glass Beads

In order to calibrate any one sieve, it is merely necessary to place the entire sample on the sieve, shake until the rate of passage of beads through the sieve is practically zero,[8] and then carefully weigh the beads that have passed the sieve and calculate the percentage by weight. The effective opening of the sieve is then read directly from the calibration curve. When several sieves are to be calibrated at the same time, they can be nested and shaken together. The weight finer than any particular sieve is the sum of the weight on all the sieves below and in the pan.

A numerical example of the procedure may be instructive. A single sieve (No. 45) was tested with a glass-bead mixture weighing 110.06 g. It was found that 25.96 g passed the sieve. Hence, the percentage by weight passing was $(25.96/110.06)\times 100=23.6$ percent. From the calibration curve for this sample (see fig. 1) the effective sieve opening is read as 372μ. The nominal opening of a No. 45 sieve is 350μ, with an allowable variation of ± 5 percent. The effective opening of this sieve is 6.3 percent larger than the nominal size. As will be explained later, such an

[7] Suitable beads have been obtained from the Cataphote Corp., Toledo, Ohio; Potter Bros. Inc. Ozone Park, N. Y.; Minnesota Mining & Manufacturing Co., Saint Paul, Minn.

[8] For details see section III, 2.

UNIVERSITY OF MICHIGAN LIBRARIES

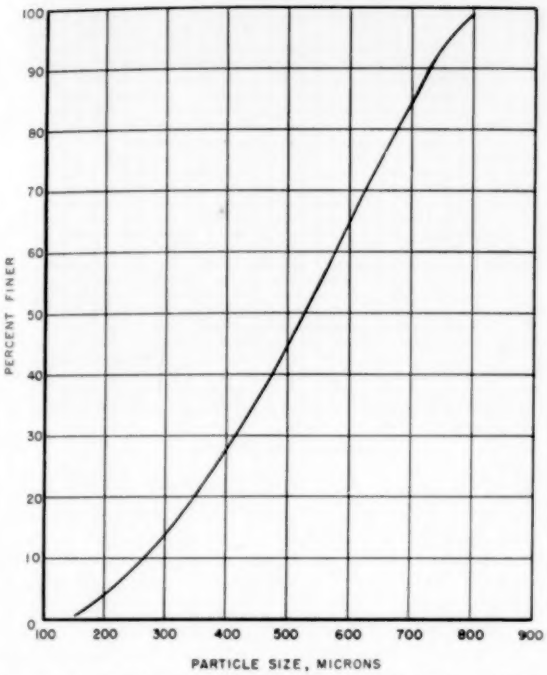

FIGURE 1. *Particle size distribution of glass beads for calibration of testing sieves.*

effect might reasonably be expected; in using this sieve for testing purposes, it is desirable to take into account the variation of its effective opening from the nominal opening.

5. Application of Correct Sieve Openings to Sieve Analyses

One of the most disturbing features of sieve analyses is the inability to obtain the same results with the same sample when using different sets of sieves. It has been pointed out that the small differences in the sieves inherent in manufacturing is the cause of this trouble, and that with suitable calibration corrections can be made for these differences. Also, if sieves that are not of standard sizes are used, the results of the sieve analysis can be reduced to what would be obtained if nominal-size sieves were used.

A simple procedure is to plot the results of sieve analysis as cumulative percentage finer as a function of the effective opening of the sieve. From this curve the cumulative percentage finer that would pass through the nominal openings can be obtained and, hence, the corrected sieve analysis may be calculated. An example of this procedure will now be given. A sample of service bone char was sieved with two different sets of sieves. Both were good sieves and supposed to be equivalent. The results are contained in table 1.

Both of these sets of sieves were then calibrated by use of the glass beads with the results given in table 2. It is to be noted that in every case the calibrated

TABLE 1. *Sieve analysis of a service bone char with two different sets of sieves*

U. S. Standard Sieve No.	Char retained on—		Ratio of percentage retained on set 1 to that on set 2
	Set 1	Set 2	
	Percent	*Percent*	
25	19.3	17.4	1.11
30	5.2	7.6	0.68
35	7.2	7.2	1.00
40	10.5	7.0	1.50
45	11.3	12.6	0.90
50	15.7	13.7	1.15
60	14.0	18.0	0.78
70	10.9	8.8	1.24
80	2.5	4.8	0.52
Through 80	3.4	2.9	1.17

opening is larger than the nominal opening. This is because this method of calibration measures an effective opening that is larger than the arithmetic average of the projections of all the openings. Moreover, the opening is effective in three dimensions, and the plane defined by the sieve cloth may not coincide with the plane defined by the effective opening. The effective opening will thus be always larger than its projection on the plane of the sieve cloth. The sieves were manufactured to have the average opening equal to the nominal opening within the specified tolerances.

The two sieve analyses are plotted as a function of their effective openings in figure 2. The two curves very nearly coincide. From these curves the

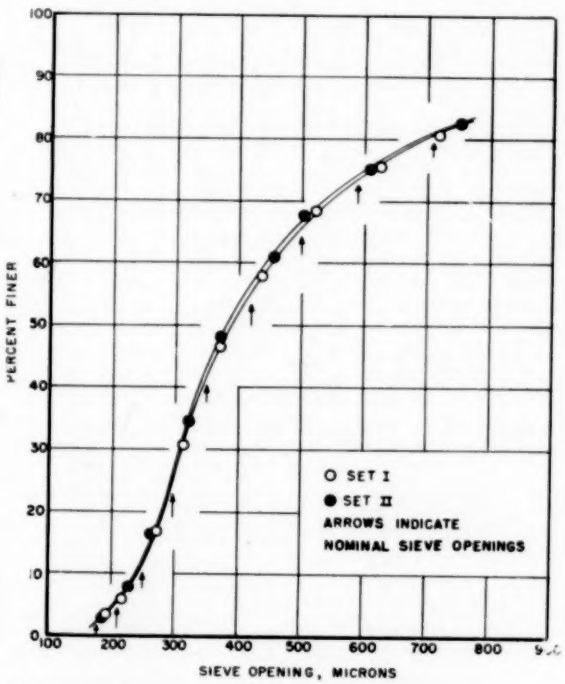

FIGURE 2. *Sieve analyses of the same sample with two sets of sieves.*

TABLE 2. *Effective openings of two sets of sieves*

[1,000μ equal 1 mm.]

U. S. Standard Sieve No.	Effective openings		Nominal opening
	Set 1	Set 2	
	μ	μ	μ
25	719	752	710
30	627	610	590
35	523	503	500
40	436	456	420
45	372	374	350
50	314	322	297
60	272	264	250
70	217	228	210
80	193	186	177

corrected analyses are read at the nominal openings (indicated by arrows). The corrected analyses are tabulated in table 3. The excellent agreement of the corrected sieve analyses indicates the value of the calibration method.

TABLE 3. *Corrected sieve analyses with different sets of sieves*

U. S. Standard Sieve No.	Corrected amount retained on—		Ratio of percentage retained on set 1 to that on set 2
	Set 1	Set 2	
	Percent	*Percent*	
25	19.7	19.4	1.02
30	7.2	7.0	1.03
35	7.1	6.8	1.04
40	10.6	10.4	1.02
45	14.0	14.1	0.99
50	16.5	15.9	1.04
60	12.6	14.0	0.90
70	7.0	7.1	0.99
80	3.4	3.2	1.06
Through 80	1.9	2.1	1.10

6. Uniformity of Openings

Besides the average size of the openings, the uniformity of openings plays an important role in sieve analysis. Before sieves are calibrated for the average effective opening they should be checked for uniformity. In figures 3 and 4 are seen two No. 10 sieves. At first glance the two look alike, but the wires in figure 3 are a little thinner than in figure 4. Both wire diameters are within specifications and, as previously stated, this difference is of minor importance. More careful examination of figure 3 reveals that the wires are not perfectly parallel, and that there is considerable variation in the size of the openings. In figure 4 the variation in the openings is scarcely discernible without the aid of measuring instruments. The sieve of figure 3 has been used for 19 years and has been subjected to rough treatment. The sieve shown in figure 4 has been used very little and still conforms to specifications.

If there are many oversized openings, as in figure 3, then many particles that should remain on the sieve pass through these openings, and the effective size of the sieve is not its average opening but more nearly its maximum opening. If only a few of the openings are very much oversize, material continues to pass these few maximum openings for a long time.

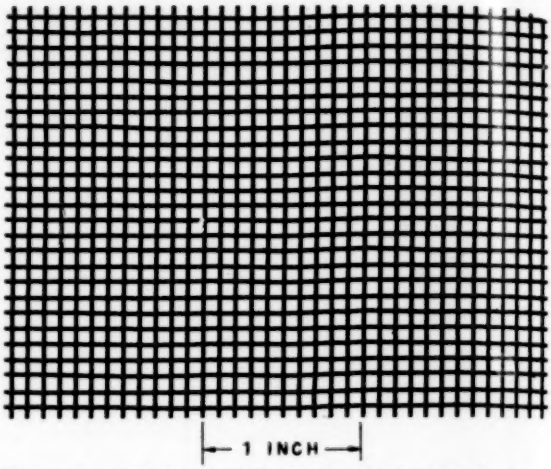

FIGURE 3. *U. S. Standard No. 10 Sieve that is visibly deformed.*

The rate of passage of material through a uniform sieve drops to a very low value after a few minutes of shaking, whereas material continues almost indefinitely to pass the few oversize openings of a nonuniform sieve.

The rate of passage of certain materials through a sieve can be taken as a measure of uniformity of the sieve. In figure 5 are seen sieving curves [9] for five different sieves. All sieves were nominally No. 40, and the same sample of glass beads was used for each. It must be remembered that it is the sieving rate (slope) rather than the actual weight passing that is of primary interest. Curve *A* of figure 5 was obtained from a badly distorted sieve whose meshes were in a condition similar to those shown in figure 3. Curve *B* was obtained from a badly worn sieve that was no longer used for testing. Curve *C* was obtained from a sieve that had been in

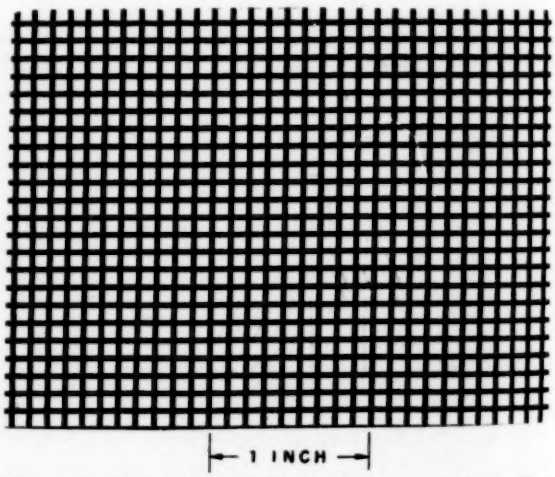

FIGURE 4. *U. S. Standard No. 10 Sieve that is in excellent condition.*

[9] Explained in detail in section III, 1.

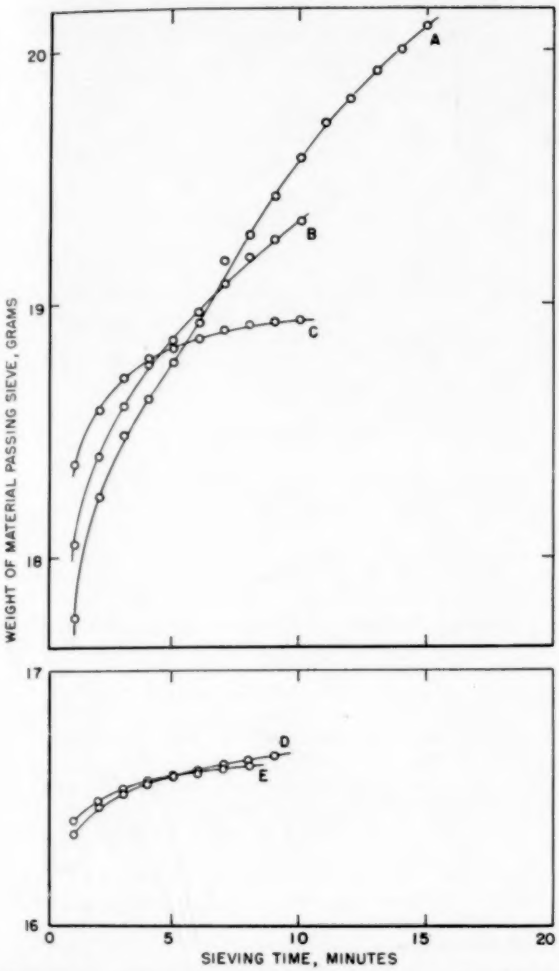

FIGURE 5. *Sieving curves (weight passing as a function of time) for sieves of varying uniformity with the same sample of glass beads.*

All sieves were No. 40.

use for some years. Curves D and E were obtained from new sieves that were certified by this Bureau to be within one-half the uniformity tolerances allowed in specifications for testing sieves. The sieves corresponding to curves, C, D, and E are considered sufficiently uniform to be of value as testing sieves, whereas those corresponding to curves A and B cannot be expected to give accurate analyses.

III. Method of Shaking Testing Sieves

The shaking methods considered were by hand and by use of the Ro-Tap machine. The standard sieving methods of the American Society for Testing Materials (ASTM) for roofing materials [9, 10] and plastic molding powders [11] prescribe the machine method as standard. The ASTM methods for testing cement [12], soap [13], powdered coal [14], refractories [15], road materials [16], and fine and coarse aggregates used in concrete [17] all prescribe hand sieving as standard, but machine sieving is acceptable for all except cement [12], provided the results agree with hand sieving. In hand sieving, the sieves are shaken one at a time until the rate of passage of material through the sieve decreases to some very low value [10] prescribed by the specifications for the particular test. Hand sieving is time consuming and tedious and, moreover, only one sieve is shaken at a time. In the machine method a stack of sieves is shaken in one operation. As it is not practical to measure the rate of passage of material through the individual sieves, the stack is shaken for a period of time long enough to insure that each particle has found its proper place.

It has been found that for some materials the agreement between hand and machine sieving is very good, whereas for other materials it is very poor. There is considerable doubt about the absolute accuracy of either method, but hand sieving is usually recommended because it is more reproducible. Machine sieving is much less time consuming and easier and, hence, if it were as reproducible as hand sieving, it would be preferred.

1. Rate of Sieving

The progress of sieving can be followed by weighing the material remaining on the sieves after convenient measured intervals of time, usually 1 to 2 min. The weight of material on the sieves is thus determined as a function of time. A graphical representation of this relationship is logically called a "sieving curve" and is illustrated in figure 6 for the different sieves in a stack.

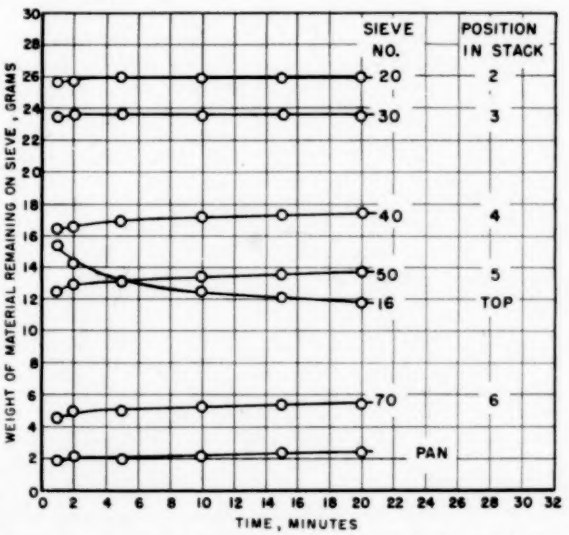

FIGURE 6. *Typical sieving curves (weight retained as a function of time) for bone char 32.*

[10] The stopping point for sieving is explained in detail in section III, 2.

At the start all the material is placed on the top sieve, and this sieve can only lose weight as the sieving progresses. The weight of material on all the other sieves is determined by the rate at which material falls upon it from above and leaves it by passing through. Each of these rates is called a sieving rate. The weight on the intermediate sieves sometimes decreases and sometimes increases, depending upon which of these sieving rates is larger. If they are approximately the same for each sieve, then the weight remaining on intermediate sieves remains constant and only the top sieve and pan will show any change, as may be seen in figure 6. Thus, the rate of change observed on the top sieve and in the pan are true sieving rates, whereas the rates of change on intermediate sieves are differences between sieving rates.

In order to study sieving rates it is quite obvious that, at most, only two sieves can be shaken at a time. The material on the top sieve and the material in the pan are weighed every few minutes to determine the two sieving rates. For practical purposes it is far better to shake only one sieve at a time. The sieving rate can be obtained from the weight of material falling into the pan, and in this way the material on the sieve is not disturbed during the weighing.

A typical sieving-rate curve is shown in figure 7. It is seen that the rate drops off very rapidly at first, but then much more slowly, and was never observed to reach zero. This is further demonstrated in table 4, in which the shaking was carried on for 80 min,

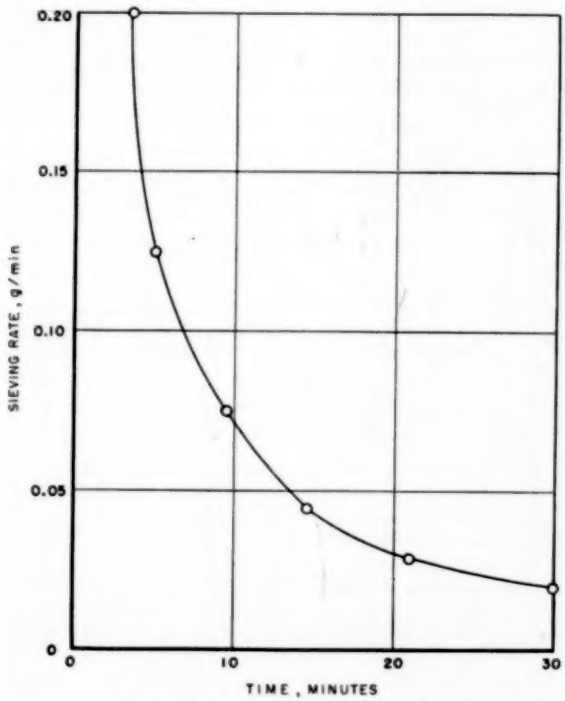

FIGURE 7. *Typical sieving rate curve for bone char.*

TABLE 4. *Effect of prolonged shaking on the Ro-T p*

Time of shaking	Weight on sieve	Time of shaking	Weight on sieve
min	*g*	*min*	*g*
2	19.0	15	16.3
3	18.3	20	16.0
4	17.8	36	15.1
5	17.7	50	14.6
7	17.0	60	14.5
9	16.8	70	14.3
11	16.7	80	14.0
13	16.4		

and the weight of bone char on the sieve was still decreasing. The continued passage of material after such long periods is due to attrition of the particles. Because of this attrition, it is desirable to shake the sieves for a short time only.

2. End Point of Sieving

The end point of sieving would be attained when no more material passed any of the sieves in the stack, or when all sieving rates became zero. However, as has been shown, the sieving rate never reaches zero, and therefore a definite end point of sieving is not attainable. It is necessary, therefore, to employ another well-defined and reproducible point that may serve as a close approximation. A designated small value of the sieving rate serves this purpose. In order that the end point be reproducible the sieving rate must be very low, so that if shaking is carried on a minute or so too much or too little, the change in the quantity on the sieve is insignificant. The magnitude of an insignificant change depends upon the required accuracy. A few tenths of 1 percent of the weight of the sample is sufficient reproducibility for many purposes.

The standard sieving methods of the ASTM for cement [12], powdered coal [14], and paving materials [16] require that the terminal rate be 0.05 g/min, and for refractories [15], and soap [13] a terminal rate of 0.1 g/min. For nongranular roofing materials [10], the rate is 0.05 percent of the weight of the total sample per minute, whereas for fine and coarse aggregates used in concrete [17] it is 1 percent of the residue on the sieve per minute. A terminal rate expressed as a percentage of the sample insures that the same accuracy is obtained for a sample of any size. However, since it can be shown that sieving rate is nearly independent of weight of sample, the terminal rate is better expressed in terms of the sieving rates.

A suitable stopping point for bone char was found to be a sieving rate of 0.1 g/min through a standard 8-in.-diameter sieve. A lower value was not suitable because of the abrasion of some soft chars. For this rate to be of the order of magnitude of 0.1 percent/min, a sample of about 100 g should be used. The same stopping point was used for both hand and machine sieving. When shaking by hand, the sieves were tapped and shaken about 100 to 130 times per min, and as the end point was approached, the amount passing was weighed each minute.

TABLE 5. *Reproducibility of hand sieving during repeated analyses of the same sample of char*

Sieve fraction	Repeated sieve analyses			Mean	Standard deviation from the mean
	1	2	3		
	CHAR 69 B GRAB				
	σ_c	σ_c	σ_c	σ_c	σ_c
On 12	2.3	2.4	2.3	2.3	0.071
12 to 16	12.8	12.9	12.8	12.8	.071
16 to 20	22.7	22.6	22.9	22.7	.158
20 to 30	25.1	25.3	25.1	25.2	.123
30 to 40	18.7	18.6	18.6	18.6	.071
40 to 50	12.4	12.4	12.3	12.4	.071
Through 50	6.0	5.8	6.0	5.9	.123
				Avg...	0.098
	CHAR 32 GRAB				
On 16	4.3	4.4	4.4	4.4	0.071
16 to 20	20.9	20.9	20.8	20.9	.071
20 to 30	25.3	25.2	25.0	25.2	.158
30 to 40	20.0	19.6	20.0	19.9	.235
40 to 50	17.3	17.5	17.5	17.4	.123
50 to 70	8.0	8.1	8.0	8.0	.071
Through 70	4.2	4.3	4.3	4.3	.071
				Avg...	0.114

Several repeated analyses with hand sieving are given in table 5. It is seen that the standard deviation from the mean is approximately 0.1 percent of the weight of the total sample.

The time required to reach the terminal sieving rate of 0.1 g/min on a Ro-Tap machine is called the sieving time. It is determined by the point at which the slope of the sieving curve becomes 0.1 g/min.

Fagerholt [2] derived a purely theoretical mathematical expression for the relation between the weight (W) of material remaining on a sieve and the time of shaking (t).

$$W = W_\infty + \frac{C}{\sqrt{t}},$$

where W_∞ is the weight that would remain after an infinite time, and C is a constant.

In his derivation it was necessary for Fagerholt to assume that (1) the rate of passage is proportional to the weight that can pass the sieve, (2) probability of passing depends upon size of particle relative to size of opening, (3) particle size distribution is constant in the range near the sieve opening, (4) sufficient time has elapsed for the passage of particles very much smaller than the sieve opening, and (5) the sieve openings are absolutely uniform in size.

It is shown in section IV, 5, that Fagerholt's second assumption is not strictly correct, because the probability of passing is also dependent upon the relative motion of particle and opening. The fifth assumption is seldom realized even for new sieves. The nonuniformity may be very great for old, distorted sieves (see fig. 3).

The rate of passage of material through the sieve is obtained by taking the derivative of the above expression.

$$\frac{dW}{dt} = \frac{-C}{2} t^{-1.5}.$$

When the rate of passage of material through a sieve is plotted against time of shaking on log paper, a straight or very slightly curved line results (fig. 12) with a slope between -1.1 and -1.7. The failure to achieve always a slope of -1.5 can be attributed to nonconformity to the assumptions used in the derivation, especially the second and fifth. Nevertheless, the plot does serve as a convenient means to determine the sieving time.

3. Ro-Tap Machines

There are two types of Ro-Tap machines in use today. The older model has two eccentrics, whereas the model being produced today has an eccentric on one side and a reciprocating motion on the other. The older model thus gives a rotary motion in the horizontal plane, and the newer model gives a combined rotary and reciprocating motion. In both machines the ratio of rotations to taps is equal to 1.875. The substance upon which the knocker strikes can be varied, and commonly used materials are cork, rubber, and hardwood.

To obtain some idea of the performance of various Ro-Tap machines, samples of two different service chars were sieved in seven different Ro-Taps and by hand. The same samples were sieved repeatedly to eliminate any sampling error. The magnitude of the attrition was found to be negligible by making the first and last analysis on the same Ro-Tap machine under identical conditions (see table 6). The same set of sieves was used by the same operator in all cases. The shaking was continued for 10 min for all tests except hand sieving. The Ro-Tap machines tested were in daily use, testing various materials and were used as found without alteration in any way. The characteristics of the various machines and the sieve analyses are recorded in table 6. The results of machine No. 1 are in closest agreement with hand shaking in these cases. However, this is not always the case, as will be shown later in the section on speed of Ro-Tap.

The coefficient of variation for the top sieve was found to be 23 percent, whereas for all others it ranged from 2 to 6 percent. This is because the weight on the top sieve can only decrease, whereas on all other sieves material falls through about as fast as it falls upon the sieves. The weight in the pan can only increase. Thus, if one shaking method causes more particles to pass the sieves than another, the top sieve and pan should indicate the greatest difference with very little difference on the intermediate sieves. An examination of table 6 shows that this is what happens, except that a large variation did not appear in the pan. Apparently, the variation that should have appeared in the pan was absorbed by the intermediate sieves.

4. Variation Due to Ro-Tap Knockers and Position of Sieves in Stack

It was decided to test the two types of Ro-Tap machines having one or two eccentric bearings with various knocker cushions and various chars on each.

TABLE 6. *Variation among Ro-Tap machines*

Machine	1	2	3	4	5	6	7	• 1	Hand	Mean of 7 machines, m	Standard deviation from mean, σ	Coefficient of variation, σ/m
Speed (taps/min)	110	155	155	160	160	158	156	112				
Eccentrics	2	2	1	1	2	1	1	2				
Knocker	Rubber	Rubber	Cork	Rubber	Rubber	Maple	Cork	Rubber				
CHAR 32												
On sieve No.:	%	%	%	%	%	%	%	%	%	%	%	%
16	4.9	5.6	5.6	5.9	8.6	4.6	5.2	4.7	4.3	5.77	1.325	23.0
20	21.6	22.9	22.5	22.4	22.3	22.3	22.4	21.4	20.9	22.34	.387	1.7
30	25.0	25.1	24.2	24.2	24.0	24.8	24.8	25.0	25.3	24.59	.441	1.8
40	19.7	20.0	19.4	19.3	19.2	19.5	19.4	19.8	20.0	19.50	.271	1.4
50	16.8	15.3	16.5	16.2	15.1	16.8	16.5	17.0	17.3	16.17	.697	4.3
70	7.8	7.4	7.7	7.8	7.1	7.7	7.6	7.8	8.0	7.59	.255	3.4
Pan	4.2	3.7	4.1	4.2	3.7	4.3	4.1	4.3	4.2	4.04	.244	6.0
CHAR 69 B												
12	2.5	3.9	2.8	3.4	4.5	2.7	2.7	2.3	2.3	3.21	0.750	23.4
16	13.5	14.7	14.0	14.2	15.4	13.6	14.0	13.1	12.7	14.20	.661	4.6
20	22.7	22.0	22.7	22.5	22.4	23.0	22.6	22.9	22.9	22.56	.310	1.4
30	25.3	25.2	24.6	24.5	24.4	24.7	24.9	25.4	25.1	24.80	.347	1.4
40	18.4	17.9	18.2	18.0	17.5	18.3	18.2	18.4	18.6	18.07	.304	1.7
50	12.0	11.0	12.1	11.8	10.7	12.0	12.0	12.2	12.3	11.66	.565	4.8
Pan	5.6	5.3	5.6	5.6	5.1	5.7	5.6	5.8	6.0	5.50	.216	3.9

• Included as a check for attrition.

TABLE 7. *Weight of material remaining on sieves for the 16 experimental variations*

	Knocker K_1	Knocker K_2	Knocker K_3	Knocker K_4
TOP SIEVE				
	g	g	g	g
Machine 1 with char 1	21.1	21.0	20.8	22.8
Machine 2 with char 2	15.8	15.9	15.6	16.0
Machine 2 with char 1	21.4	21.4	21.1	21.5
Machine 1 with char 2	15.0	14.9	15.2	16.9
SECOND SIEVE				
Machine 1 with char 1	17.0	17.3	17.1	17.7
Machine 2 with char 2	18.7	18.6	18.6	18.6
Machine 2 with char 1	17.2	17.3	17.4	17.3
Machine 1 with char 2	18.6	18.7	18.6	18.4
THIRD SIEVE				
Machine 1 with char 1	16.6	16.1	16.6	15.4
Machine 2 with char 2	14.7	14.7	14.8	14.6
Machine 2 with char 1	16.1	16.0	16.1	16.0
Machine 1 with char 2	14.9	14.9	14.7	15.3
PAN				
Machine 1 with char 1	16.2	16.2	16.2	14.8
Machine 2 with char 2	15.7	15.6	15.8	15.5
Machine 2 with char 1	16.0	15.9	16.1	15.9
Machine 1 with char 2	16.3	16.3	16.3	14.2

The experiments were arranged to give the greatest amount of information by application of statistical methods of evaluation.[11] Two different service chars were used on two different machines. The use of three different knocker cushions and the omission of the knocker made four different variations of knockers. Three sieves and a pan were used. The time for each shaking was 7.5 min. A tabulation of the variations employed is given below:

Machines:

M_1—2 eccentrics at 115 taps/min.
M_2—1 eccentric at 150 taps/min.

Chars:

C_1—70.7 g of char 32 with sieves No. 18, 25, 35, pan.
C_2—64.8 g of char 33 with sieves No. 20, 30, 40, pan.

Knocker cushions:

K_1—Rubber.
K_2—Cork.
K_3—Hardwood (maple).
K_4—None.

Experiments were undertaken with combinations of each of the possible machines, chars, or knocker cushions by using the sieves designated for each char. To eliminate any possible bias due to attrition of the char, the order of the experiments was taken from a Latin square arrangement. The Latin square is a statistical method of removing bias due to the order of experiment, as shown here.

Order of experiments

	Knocker K_1	Knocker K_2	Knocker K_3	Knocker K_4
Machine 1 with char 1	1	2	3	4
Machine 2 with char 2	6	5	8	7
Machine 2 with char 1	11	12	9	10
Machine 1 with char 2	16	15	14	13

The results for each sieve are given in table 7.

From an analysis of variance (see appendix), it was found that the variation due to knockers, chars, and the interaction between machines and knockers

[11] The authors are indebted to John Mandel for aid in the design of the experiments and their evaluation by statistical methods.

UNIVERSITY OF MICHIGAN LIBRARIES
DEPARTMENT OF MICHIGANA COLLECTIONS

is highly significant and the interaction between machines and chars is significant. Table 7 shows that the absence of a knocker (condition knocker 4) is the greatest cause of the variation. This means that sieve analyses obtained without the knocker are not comparable to those in which a knocker is used. When the analysis of variance is repeated omitting K_4 (see appendix), it is found that there is significant variation between chars on all sieves and that there is significant variation between machines in the case of the top sieve and pan. There is no significant variation in any other case.

That a difference was found between chars was to be expected. The important point is that there is a difference between machines but none among knockers nor any interactions among combinations of knockers, chars, or machines. It was found that the magnitude of the difference between machines was dependent upon the speed of the mechanism, and this point is covered further under that heading. That no significant difference was found among knockers indicates that this is not an important source of error in machine sieving and need not be rigidly controlled. The fact that the interactions with chars were not significant indicates that all chars react in the same way and that a sieving procedure that will produce good results for one char will also produce good results for other chars.

Some of the older machines have flat leaf springs to hold the sieves at the bottom instead of the cast iron plate, as in the newer models. It was found that the tension in these springs had no effect upon the sieving.

When materials of different grist are sieved, it is expedient to have a certain sieve sometimes at the top of the stack and sometimes at the bottom. To determine whether or not a sieve gives the same results in the various positions, samples of service char were sieved on a single sieve at various positions in the stack. All the other sieves that filled up the stack were very much larger, and no material was retained on any of them. It was concluded that the position of the sieve in the stack is of no importance, as far as the results on that one sieve are concerned.

5. Speed of Ro-Tap

Early experiments showed that the speed of the Ro-Tap has an important effect on the final weight of material that remains on a sieve, and that this effect depends upon the size of opening. Therefore, a series of experiments was undertaken in which sieving curves were determined for closely sized fractions of char at various speeds of the Ro-Tap. A weight of 15.00 g of each of the following char fractions was placed on the sieve designated below, and the sieving repeated at various speeds. A fraction of char previously sieved to pass No. 12 and be retained on No. 16 was sieved on No. 14; similarly, a 30 to 40 fraction was sieved on a No. 35; and a 70 to 100 fraction on sieve No. 80. The general shape of the sieving curves is the same for all sizes, and typical results are shown in figure 8. It is noted that the curve at the high speed (150 taps/min) levels

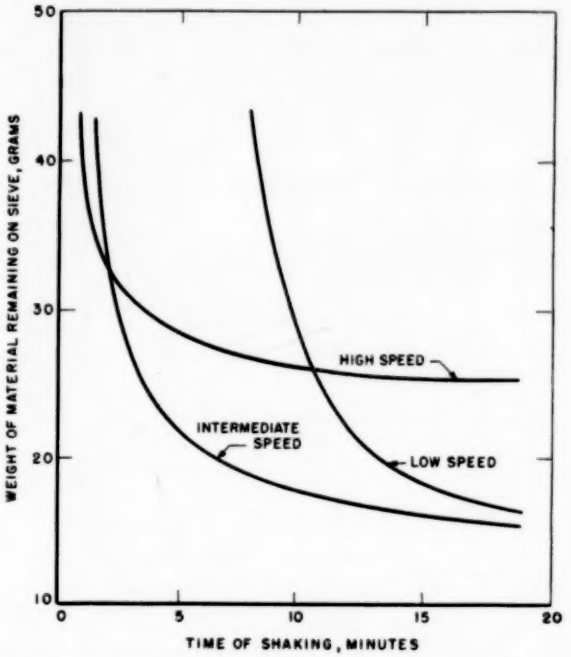

FIGURE 8. *Effect of Ro-Tap speed on weight retained as a function of time.*

off much higher than the curve for intermediate speeds (115 taps/min), whereas the curve at low speeds (less than 90 taps/min) levels off at the same value as that for intermediate speeds but takes more time to do it.

This effect is brought out more clearly in figures 9 and 10. The weight of material on a sieve after 10 min of shaking is plotted as a function of speed in figure 9. In figure 10 the weight of material remaining on the sieve when the sieving rate falls to 0.1 g/min (termination of sieving) is plotted as a function of speed and compared to hand shaking. In figure 9 it is seen that there is a definite minimum in the amount remaining on the sieves at a speed of about 115 taps/min; the curves in figure 10 merely flatten out at low speeds. This means that at high speeds (150 taps/min) the sieve retains some particles that pass through it at lower speeds. At 115 taps/min the minimum amount is held on the sieve, and further reduction of speed only serves to lengthen the time required to complete the sieving.

Since it is desirable to have the sieve analysis give the closest possible representation of particle size distribution, it follows that the most accurate sieve analysis will be obtained at the speed at which the minimum occurs, namely 115 taps/min. Also, sieve analyses made at the speed of the minimum should be more reproducible, because adjacent to the minimum the curves slope very slightly. Thus, if the speed of Ro-Tap varies by a few taps a minute, then the effect on the amount passing the sieve is minimized. At the usual speed of 150 taps/min, the

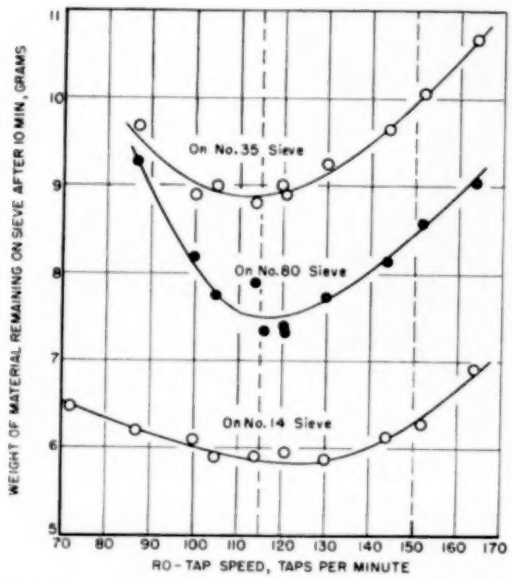
WEIGHT OF MATERIAL REMAINING ON SIEVE AFTER 10 MIN, GRAMS
On No. 35 Sieve
On No. 80 Sieve
On No. 14 Sieve
RO-TAP SPEED, TAPS PER MINUTE

FIGURE 9. *Effect of Ro-Tap speed on the weight of bone char retained on various sieves after 10 minutes.*

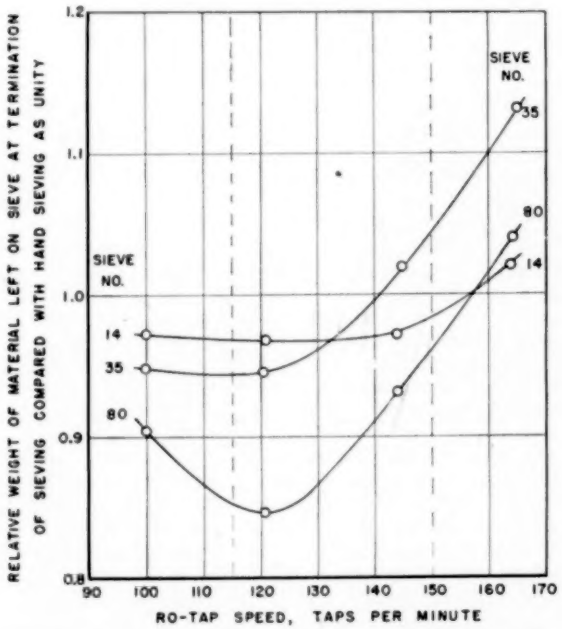

FIGURE 10. *Effect of Ro-Tap speed on the weight retained on various sieves at the termination of sieving compared with hand sieving as unity.*

slopes of the curves in figures 9 and 10 are such that a variation of 5 taps/min changes the amount remaining on the sieve by about 2 percent. The variations noted in table 6 are largely due to variations in speed. The reproducibility of sieve analyses made at one constant speed is essentially the same for all types of machines regardless of the speed.

The agreement between hand and machine sieving at a speed of 150 taps/min is somewhat better than at a speed of 115 taps/min. However, the improved reproducibility and the apparent greater accuracy of machine analyses at 115 taps/min more than outweighs the disadvantage of nonagreement with hand sieving for materials such as bone char.

To study the agreement of the two types of Ro-Tap machines at various speeds, three samples of char having different particle sizes were run on both types. The weights on the sieves after 10 min of sieving are shown as a function of speed in figure 11. The results are in substantial agreement for all particle sizes at speeds below 115 taps/min. At higher speeds the results obtained on the two machines become more divergent. The two-eccentric machine, in which the motion is more violent, does not pass as much material as the one-eccentric machine per unit of time.

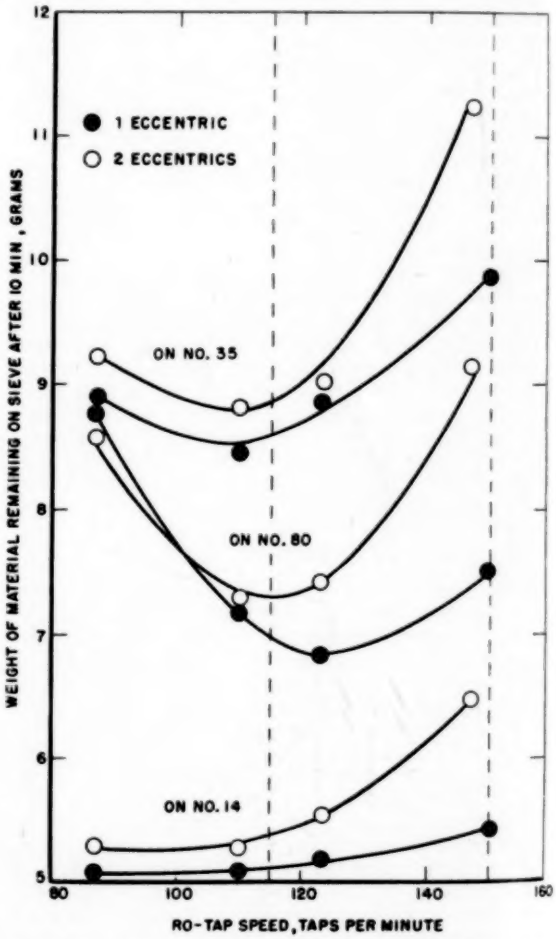
1 ECCENTRIC
2 ECCENTRICS
ON NO. 35
ON NO. 80
ON NO. 14
WEIGHT OF MATERIAL REMAINING ON SIEVE AFTER 10 MIN, GRAMS
RO-TAP SPEED, TAPS PER MINUTE

FIGURE 11. *Effect of Ro-Tap speed with two types of machines.*

Fahrenwald and Stockdale [19] studied the effect of the motion of the sieve. They found that the highest rates of transport of material through the sieve could be obtained at very high speeds of shaking if the amplitude of motion was about the same as the size of the sieve openings. The amplitude of motion of the Ro-Tap machine is about 1 in., which is much larger than the opening of any sieve studied here. Apparently the selective retention of *slightly undersized particles* at high speeds of the Ro-Tap is to be attributed to excessive amplitude of motion in the horizontal plane. The enhanced retention of *all particle sizes* at low speeds is due to the lack of sufficient motion. It is possible to express the probability, P, of a particle passing through a sieve opening in terms of fundamental mechanics. Each particle can be uniquely located in the space above the sieve by six coordinates, three of translation and three of momentum. The probability, P, is then a ratio, in which the denominator is the product of all possible positional and momentum values in phase space, and the numerator is the particular positional and momentum values required for the particle to pass the opening. All types of sieving devices that utilize forced vibrations should be subject to this analysis and exhibit an optimum value of frequency for maximum transport through the sieve openings.

6. Effect of Speed of Ro-Tap for Various Materials

In order to determine whether the effect of variations in speed was peculiar to bone char or general for all types of material, a number of different materials, particle shapes, and particle sizes were examined. The results are summarized in table 8. The weight of material remaining on the sieves varied over wide ranges, but since the actual weight is not important the amounts remaining on the sieves are expressed relative to the minimum amount. The data should be regarded as qualitative. The fact that one material shows a 10-percent and another a 50-percent increase in the amount retained on the sieve at high speed is of no significance; it only reflects the different particle size distributions. The same behavior with regard to speed was found for all the materials tested; that is, at a speed of Ro-Tap of approximately 150 taps/min some under-sized particles are retained that will pass through at a speed of approximately 115 taps/min.

TABLE 8. *Effect of Ro-Tap speed for various materials*

Material	Particle shape	U. S. Standard Sieve No.	Relative amount on sieve	Speed
				Taps/min
Glass	Beads	25	1.08	88
			1.00	116
			1.08	150
Do	do	6	1.37	91
			1.00	120
			1.38	152
Char 32	Granular	50	1.00	118
			1.09	155
Char 68 C	do	8	1.00	122
			1.52	158
Bismuth	do	30	1.00	122
			1.04	156
Iron	Filings	30	1.00	122
			1.08	156
Sand	Granular	25	1.09	117
			1.25	156
Floridin (73 H)	Rounded	40	1.00	117
			1.13	156

IV. Sample for Sieve Analysis

It is of importance that the sample for sieve analysis be representative of the material from which it was taken. However, this study is primarily concerned with the influence of weight and other physical properties of the sample on the sieving. The sample for sieve analysis is usually only a small fraction of the original material and must be prepared by coning and quartering or the use of a riffle, or the use of other suitable sample divider. The entire end product of the sample-reduction process must be used as the sample for the sieve analysis, otherwise it would completely defeat the purpose of the reduction process.

1. Weight of Sample

The minimum number of particles that can be considered a representative sample of a heterogeneous material such as bone char is about 10,000, and it would be preferable to have many more. From this point of view, the sample for sieve analysis should be large enough to have this number in each fraction. Ten thousand particles of 12 to 14 fraction of bone char weigh about 30 g, and 10,000 particles of 80 to 100 fraction weigh about 30 mg.

The testing sieves are only 8 in. in diameter, and if they are loaded too heavily a condition known as "blinding" occurs. When this takes place, nearly all the openings become plugged by particles wedged into them, and even material very much finer than the sieve opening cannot pass. Too great a depth of material on the sieve and the hammering of the oversize pieces into the openings by the many pieces above contribute to the blinding. It is generally agreed that the ideal depth of material on the sieve is no more than one or two particles. A layer two particles deep of 80 to 100 fraction of bone char with a bulk density of 60 lb/ft^3 on an 8-in. sieve weighs about 10 g. A layer of 10 to 12 fraction of material this deep weighs 100 g. A layer two particles deep contains more than 10,000 particles for sieves finer than No. 8. It has been found that material four or six particles deep can be successfully sieved, but more than this causes excessive blinding.

In the range of particle sizes commonly found in granular materials, a convenient weight of sample is 100 g. Samples of bone char larger than 500 g always produce excessive blinding of the sieve, and samples smaller than 50 g always give very erratic results, because the coarser fractions contain too few particles. Since it is not practicable to prepare a sample of exactly 100.00 g by a sample reduction process, approximately 100 g must be used. The effect of sample size must be determined so that sieve analyses made with different sized samples can be reduced to a comparable basis. Preliminary ex-

periments showed that identical sieve analyses could be obtained with samples from 50 to 500 g if the shaking was continued just long enough to reach the end point of sieving.

It might be expected that the sieving time required to reach this end point would be directly proportional to the weight of sample. This is found to be true for any one particular sample. However, further investigation showed that the sieving time was also dependent upon the particle size distribution, particle size, the particle shape, and sieve uniformity.

2. Effect of Particle Size Distribution

An experiment was performed to study the influence of those particles that are of such sizes that they might hinder the passage of the particles that are a near fit. A service bone char (char 32) was carefully sieved into the following fractions:

Fraction	Weight
	g
Through No. 25 on No. 30	90
Through No. 30 on No. 35	10
Through No. 35 on No. 40	10
Through No. 40 on No. 45	90

First, the 30 to 35 and 35 to 40 fractions (total weight 20 g) were mixed together and sieved on the No. 35 sieve and the rate determined as a function of time. Then the 25 to 30 fraction was added (total weight 110 g) and the sieving repeated; finally, all four fractions were mixed (total weight 200 g) and again the sieving repeated. The results are given in table 9 and shown graphically in figure 12. It is at once apparent that the same value for sieving time was obtained for all three experiments, although the weight of the sample varies tenfold, and that many over-sized particles do not effect the passage of the near-fit particles. The near-fit particles are defined as those that pass the next larger sieve and are retained on the next smaller sieve in $\sqrt[4]{2}$ series. The effect of sieve uniformity does not enter here, because only one No. 35 sieve was used. From these experiments it can be concluded that the sieving time is not affected by the weight of the entire sample nor by the weight of that portion that remains on top of the sieve, but rather by the weight of the near-fit particles.

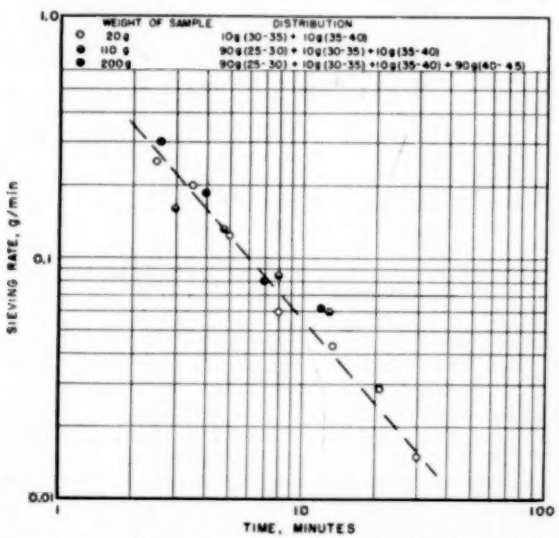

FIGURE 12. *Effect of size of sample and particle size distribution on sieving time.*

TABLE 9. *Effect of size of sample and relative particle size distribution on sieving time of fractions of char 32*

Time interval	Additional weight through	Weight rate	Mean time
10 g (30 to 35)+10 g (35 to 40) on No. 35			
min	*g*	*g/min*	*min*
2 to 3	0.25	0.25	2.5
3 to 4	.20	.20	3.5
4 to 6	.25	.125	5
6 to 10	.24	.060	8
10 to 17	.30	.043	13.5
17 to 25	.23	.029	21
25 to 35	.15	.015	30
90 g (25 to 30)+10 g (30 to 35)+10 g (35 to 40) on No. 35			
2.5 to 3.5	0.16	0.16	3
3.5 to 6	.33	.132	4.75
6 to 10	.34	.085	8
10 to 16	.36	.060	13
90 g (25 to 30)+10 g (30 to 35)+10 g (35 to 40)+90 g (40 to 45) on No. 35			
2.16 to 3	0.25	0.30	2.58
3 to 5	.37	.185	4
5 to 9	.32	.08	7
9 to 15	.37	.062	12

3. Effect of Particle Size on Sieving Time

The near-fit material has to fall upon the openings a large number of times before it is known whether or not it passes. Accordingly, the sieving time should be dependent upon the number of near-fit particles per sieve opening. The number of particles (N) is proportional to the weight of material (W) and inversely proportional to the cube of the diameter of the particle (d), ($N=k_1\,W/d^3$). The number of sieve openings (H) in a constant sieve area is inversely proportional to the square of the diameter of the opening,[12] ($H=k_2/d^2$). In these relationships k_1 and and k_2 are constants. Thus, the number of particles per sieve opening is given by the following relationship: $N/H=(k_1/k_2)\times(W/d)$. Hence, the sieving time should be dependent upon the weight of near-fit material divided by the size of the sieve opening.

To check this hypothesis, experiments were conducted with sieves of certified uniformity in which

[12] More exactly, the number of openings is proportional to the free or open area, but the wire diameters and sieve openings are such that the ratio of open area to total sieve area is about the same for all sieve sizes.

the weight of near-fit particles was varied over wide limits for several different sieve sizes and particle shapes. The sieving times are given in table 10. In figure 13 the sieving time is plotted as a function of the weight of near-fit particles divided by the size of the opening. Sizes equal to or smaller than

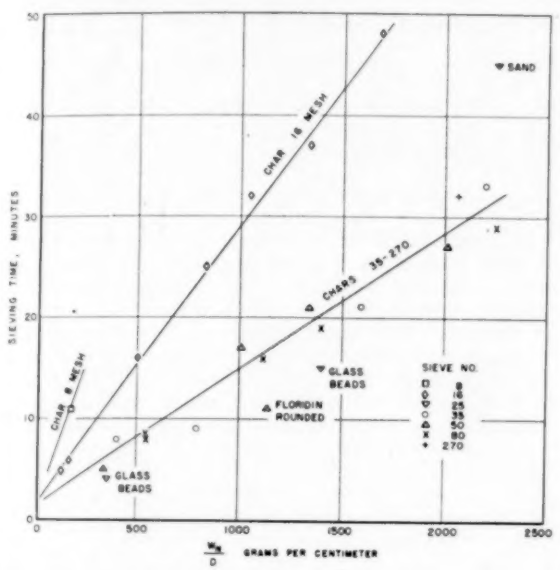

FIGURE 13. *Correlation of sieving time with the number of near-fit particles per sieve opening.*

The number of near-fit particles per sieve openings is proportional to W_N/D.

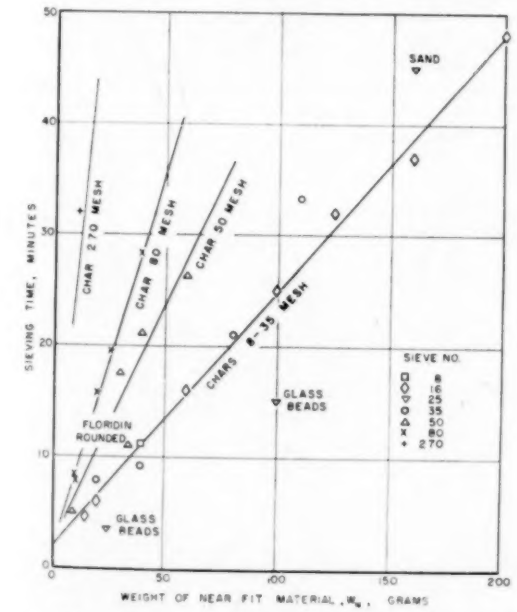

FIGURE 14. *Dependence of sieving time on weight of near-fit material.*

TABLE 10. *Sieving times for various weights, sizes, and shapes of particles*

Material	Weight of near-fit, W_N	Sieve No.	Diameter of opening d	Sieving time	W_N/d
	g		*cm*	*min*	*g/cm*
Chars	40	8	0.238	11	168
Do	500	16	.119	115	4,200
Do	200	16	.119	48	1,680
Do	160	16	.119	37	1,345
Do	125	16	.119	32	1,050
Do	100	16	.119	25	840
Do	60	16	.119	16	504
Do	20	16	.119	6	168
Do	15	16	.119	4.5	126
Do	110	35	.050	33	2,200
Do	80	35	.050	21	1,600
Do	40	35	.050	9	800
Do	20	35	.050	8	400
Do	60	50	.0297	27	2,020
Do	40	50	.0297	21	1,346
Do	30	50	.0297	17	1,010
Do	10	50	.0297	5	337
Do	40	80	.0177	29	2,260
Do	25	80	.0177	19	1,412
Do	20	80	.0177	16	1,130
Do	10	80	.0177	8	565
Do	10	80	.0177	8.5	565
Do	11	270	.0053	32	2,070
Sand (angular)	160	25	.071	45	2,250
Glass (spherical)	100	25	.071	15	1,410
Do	25	25	.071	4	350
Floridin (round)	35	50	.0297	11	1,170

the opening of No. 35 fall near the same straight line, but sizes coarser than No. 35 fall near lines of increasing slope. If these sieving times are plotted as a function of the weight of near-fit particles (fig. 14), then sizes equal to or coarser than No. 35 fall near one line and those finer on lines of increasing slope. The reason for this behavior is not apparent. It may be noted from figures 13 and 14 that the round shapes require only about one-half as much sieving time as the irregular shapes. As all curves have an intercept of 1 or 2 min for zero weight of material, this is the time required for particles tha are very much finer than the sieve to pass through.

For bone chars and other granular materials of irregular shapes, the sieving time (T, min) can be represented as a function of the weight of near-fit material (W_N, g) and the sieve opening (d, cm) by the following expressions:

For sieves No. 35 and coarser:
$$T\ (\text{min}) = 2 + 0.23\ W_N$$

For sieves No. 35 and finer:
$$T\ (\text{min}) = 1 + 0.0137\ W_N/d.$$

The times should be calculated to the nearest minute. For rounded particles, somewhat less time is required with a minimum of one-half of the values obtained from these equations. The time of shaking of a stack of sieves should be determined by the sieve requiring the maximum time. In the usual sieving operation the approximate sieve analysis is not previously known, so that there is no way of determining the sieving time until a trial run is completed. For the trial run, 10 min is a suitable shaking time for samples of about 100 g.

It has been noted that when sieving service bone chars with the $\sqrt{2}$ series of sieves, there is usually a maximum of 20 to 35 percent of the total weight (W) of the sample on any one sieve. Hence, the weight of near-fit W_N on that sieve is to a first approximation about three-tenths of the total sample, namely, $W_N = 0.3\ W$. The approximate sieving time for service chars can then be simply ascertained for a maximum occurring on a sieve coarser than No. 35:

$$T(\text{min}) = 2 + 0.07\ W,$$

and for a maximum occurring on a sieve finer than No. 35:

$$T(\text{min}) = 1 + 0.004\ \frac{W}{d}.$$

It is recommended that a sample weighing not more than 150 g nor less than 100 g be used and the sieving time be taken from the above relations, which are expressed graphically in figure 15. If a maximum of more than 35 percent or less than 20 percent be retained on any sieve, or if the sieves used are not the $\sqrt{2}$ series, or, if it is necessary to use less than 100 g or more than 150 g, then figure 15 and the approximate equations do not apply. It is then necessary to use the more exact relations given in the text. For example, assume that a preliminary sieve analysis showed that a certain char dust was about 42 percent between the No. 50 and 70 sieve and about 34 percent between the No. 70 and 100 sieve. A new sample, which weighed 108 g, was prepared on the riffle. As the sieves are in the $\sqrt{2}$ series, the weight on the sieves is approximately equal to the weight of near-fit material. Hence the weight of near-fit expected on the No. 70 sieve is $0.42 \times 108 = 45.4$ g and that on the No. 100 is $0.34 \times 108 = 36.8$ g. The opening of the No. 70 sieve is 0.021 cm and that of the No. 100 is 0.0149 cm. From the relationship given for sieves finer than No. 35, the sieving time for the No. 70 is

$$T = 1 + 0.0137\, W_N/d = 1 + 0.0137 \times (45.4/0.021) = 31 \text{ min.}$$

For the No. 100 sieve

$$T = 1 + 0.0137\ (36.8/0.0149) = 35 \text{ min.}$$

The sieving time for this sample is 35 min. The No. 100 sieve required the longest time, even though the maximum weight does not occur on this sieve.

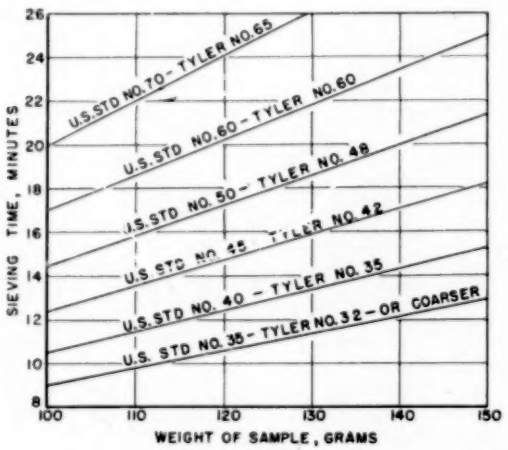

FIGURE 15. *Approximate sieving time as a function of the weight of the bone char.*

This method is to be used only if a maximum of 20 to 35 percent of the sample is retained on the sieve.

4. Test for Sieve Uniformity

Having considered the influence on the sieving rate of particle size, particle-size distribution, weight of sample, method of shaking, and shaking time, it is possible to devise a test for sieve uniformity. Any correlation between sieving rate and sieve uniformity is necessarily approximate because of the many factors involved. It has been shown that the method of shaking has practically no effect on sieving rate as long as the speed of the Ro-Tap remains constant.

The material usually employed on the sieves can be used for testing for uniformity. A sample that gives the same sieving rate for all sizes of sieves may be prepared by adjusting the weight of near-fit material for each size sieve. Five minutes was found to be a suitable sieving time. The calculated weights of near-fit material required to give a sieving time of 5 min is given in table 11 for a service bone char. The weights of individual fractions needed for a test sample of a service bone char are listed in table 12. The sum of the weights of the sieve fractions immediately above and below any given sieve is equal to the required weight of near-fit material. As it has been shown that oversized and undersized particles do not affect the sieving rate, a master test sample can be made up consisting of the tabulated amounts of all fractions in the desired range of sieve sizes. If the entire $\sqrt{2}$ series of sieves is not available, then combined fractions can be used with little difference in over-all results.

In making a test, the prepared sample is placed on one sieve and shaken for intervals of 1 or 2 mins. After each period of shaking, the additional amount in the pan is weighed and the rate of passage of material through the sieves is plotted as a function of time of

TABLE 11. *Calculated values of weight of near-fit material for a 5-minute sieving time for a service bone char*

U. S. Standard Sieve No.	Weight of near-fit	U. S. Standard Sieve No.	Weight of near-fit
	g		*g*
35 and coarser	13.04	120	3.65
40	12.26	140	3.07
45	10.22	170	2.57
50	8.67	200	2.16
60	7.30	230	1.81
70	6.13	270	1.55
80	5.17	325	1.28
100	4.35	400	1.08

TABLE 12. *Calculated weight of sieve fractions of a service bone char needed to obtain a test sample for sieve uniformity*

Sieve fraction	Weight	Sieve fraction	Weight
	g		g
30 to 35 and all coarser close fractions [a]	6.52 each	100 to 120	1.94
35 to 40	6.65	120 to 140	1.71
40 to 45	5.48	140 to 170	1.36
45 to 50	4.74	170 to 200	1.21
50 to 60	3.93	200 to 230	0.95
60 to 70	3.37	230 to 270	.86
70 to 80	2.76	270 to 325	.69
80 to 100	2.41	325 to 400	.59
		Through 400	.49

[a] Includes the fractions 4 to 5, 5 to 6, 6 to 7, 7 to 8, 8 to 10, 10 to 12, 12 to 14, 14 to 16, 16 to 18, 18 to 20, 20 to 25, 25 to 30, and 30 to 35.

shaking. If the sieves are uniform the sieving rate drops to 0.1 g/min in 5 min. or less. Typical sieving rate curves for sieves of varying uniformity are shown in figure 16. The heavy line in figure 16 has been drawn to indicate an arbitrary dividing line between satisfactory and unsatisfactory sieves. It is significant that the sieving rates for nonuniform sieves are very erratic as compared with the regular behavior of satisfactory sieves. Any closer correlation between sieve uniformity and sieving rate would require a long statistical study, which is not warranted at present.

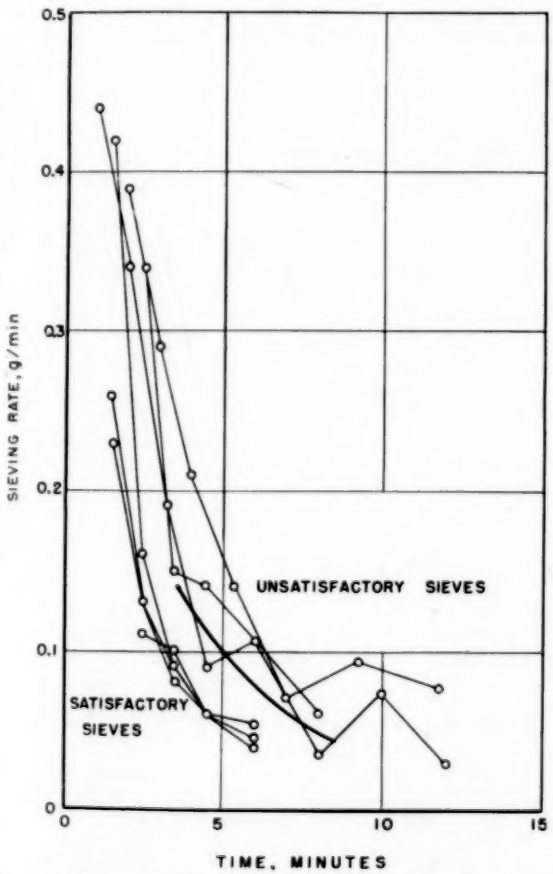

FIGURE 16. *Suggested criteria for an approximate classification of sieves with reference to uniformity of openings.*

5. Density and Hardness

Although substances ranging in density from about 1 g/cm^3 (activated charcoal) to about 10 g/cm^3 (bismuth) have been examined, no effect of particle density on sieving characteristics has been observed. The particle density for materials such as bone char and other adsorbents may be dependent upon particle size, thus complicating the relationship between the particle-size distributions by weight and by number.

As previously mentioned, material may continue indefinitely to pass through a sieve in motion because of the wearing of the particles. The minimum sieving rate, which is approached asymptotically, should be a measure of the abrasion resistance of the material. However, it would not be practical to determine abrasion resistance of bone char in this way, because it would require many hours of shaking on the Ro-Tap machine.

The resistance of the material to the type of abrasion encountered in the Ro-Tap machine can be measured very conveniently by repeating the sieve analyses several times. Any trend toward finer sizes is an indication of abrasion. In one experiment 128.6 g of a good service char was sieved five times for shaking periods of 10 min each. The weight of char remaining on the various sieves is plotted as a function of the number of sievings in figure 17. No particular trend is apparent on any sieve, but the weight in the pan definitely increases. Apparently, the wearing of bone char is such as to break off the sharp corners and thus produce fines rather than the fracture of large pieces. Attempts to detect changes on intermediate sieves fail because of the experimental error in sieving.

The rate of increase on the pan is 0.0034 g/min for the 128.6 g of sample, or 0.0026 percent/min. This rate of abrasion is quite insignificant in comparison with the other errors of sieving. The rate of abrasion of some soft chars is several times this value, and for a shaking period of 10 min, the increase in the pan fraction is sometimes as much as 0.1 percent. This is just large enough to be detected by a sieve analysis and, hence, very soft chars should be shaken as short a time as necessary.

6. Effect of Moisture on Sieve Analyses

The amount of moisture that a bone char can adsorb and still appear dry varies greatly with the char. A very good service char was found to adsorb about 17 percent of moisture (dry basis) before appearing wet, whereas a discard char adsorbed only about one-third that amount. The adsorption of water seems to be approximately proportional to the total surface area, the latter being measured by the adsorption of nitrogen at low temperature.

Small amounts of moisture have little effect upon passage of char particles through a sieve. It has been observed that bone char that appears to be dry

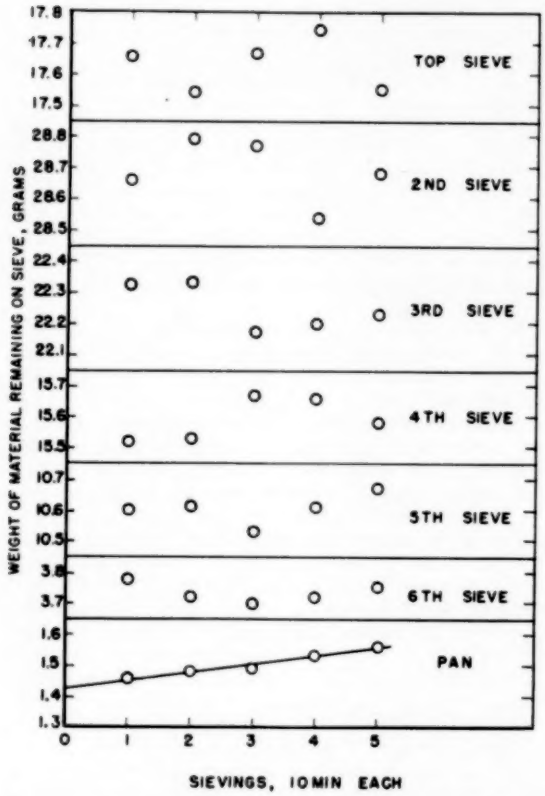

FIGURE 17. *Attrition of bone char due to repeated sievings.*
The only significant trend is the increase in the weight of the pan fraction.

sieves just like dry material. On the other hand, if a bone char appears definitely to be moist, it does not sieve. In a test run, a sample of service char appearing wet, which contained a total moisture of 17 percent, was shaken for 10 min in a nest of six sieves in the usual fashion. It was observed that more than 70 percent of the sample remained on the top sieve and none had reached the bottom three sieves.

The weights of moist char retained on the sieves will appear to be slightly different from that for the dried char, because the different particle sizes adsorb water to a different extent. Moist particles presumably would have the same size as when dry and, therefore, go through the same sieves; however, they would weigh more. Consequently, a sieve analysis of moist char must be accompanied by a moisture determination of the various fractions and the results calculated to a dry basis. The sieve analyses of a sample of chars 32 and 34 when dry and when containing different amounts of total moisture are given in table 13. A moisture determination was made for each fraction of the two wet chars and the sieve fraction corrected to correspond to dry char. When the char contains only 1 or 2 percent of moisture, the corrections are very small. Although a sieve analysis can be made with a char containing appreciable moisture, it is more practical to dry the complete sample ahead of time.

V. Summary and Recommended Procedure

It is not possible to assign a numerical value to the error to be ascribed to every particular source. It can be said, however, that the largest source of error by far is in the tolerances permitted in the average opening of testing sieves. When the standard $\sqrt{2}$ series of sieves conforming to present specifications is used, the differences when using different sets of sieves are usually in error by not more than 5 percent of the total sample. This error, due to the sieves, can be reduced to a few tenths of one percent of the total sample by determining the effective openings of the sieves with a prepared sample of spherical glass beads. However, nonuniform openings in the sieves produce erratic and inconsistent sieve analyses.

The method of shaking is a secondary source of error, in general, as long as some uniformity is observed. The position of the sieve in the stack, the knocker cushion material, and the tension of the supporting springs produce errors of less than 0.1 percent of total sample on any sieve and hence need not be rigidly controlled. It has been demonstrated that a speed of 115 taps/min is superior to the usual

TABLE 13. *Effect of moisture on sieve analysis of bone char*

U. S. Standard Sieve No.	Sieve analysis dry char	Moisture content	Sieve analysis		Moisture content	Sieve analysis		U. S. Standard Sieve No.	Sieve analysis dry char	Moisture content	Sieve analysis	
			Wet	Dry (corrected)		Wet	Dry (corrected)				Wet	Dry (corrected)
	SERVICE CHAR 32 (GRAB SAMPLE)								SERVICE CHAR 34 (GRAB SAMPLE)			
		1.28% total moisture			13.6% total moisture					4.2% total moisture		
	%	%	%	%	%	%	%		%	%	%	%
On 16	13.9	1.7	14.1	14.1	17.4	14.8	14.3	On 14	31.7	4.0	31.7	31.7
16 to 20	36.4	1.3	36.2	36.2	13.8	36.7	36.6	14 to 16	26.9	4.8	27.9	27.8
20 to 30	25.9	1.2	26.1	26.1	14.1	25.7	26.6	16 to 18	27.1	4.6	26.8	26.7
30 to 40	15.0	1.1	14.9	14.0	11.7	14.6	14.8	18 to 20	9.65	3.4	9.18	9.26
40 to 50	7.23	1.1	7.20	7.21	9.4	6.75	7.01	20 to 30	4.40	2.2	4.22	4.31
50 to 70	1.28	0.7	1.36	1.37	5.2	1.25	1.35	Through 30	0.22	2.6	0.28	0.28
Through 70	0.20		0.19	0.20		0.23	0.26					

UNIVERSITY OF MICHIGAN LIBRARIES

speed of 150 taps/min. Tne reproducibility is the same at speeds of 80 to 165, as long as the speed is constant. However, a minimum amount is retained on the sieves at 115 taps/min, and apparently the best measure of particle diameter is obtained. The two types of Ro-Tap machines (one or two eccentrics) are equivalent at 115 taps/min, but they are not equivalent at 150 taps/min. The differences between sieving analyses obtained on different machines are probably due entirely to differences in speed.

The errors inherent in the sampling and sample dividing procedure are not considered in this paper. The characteristics of the material can, however, be the source of some errors in sieve analysis. The density of the material being sieved has apparently no effect on the sieving characteristics. Nearly all bone chars are hard enough so that negligible wear occurs during a sieve analysis. Moist bone char can be satisfactorily sieved provided it does not appear moist and that the variation of moisture with particle size is properly evaluated.

The errors introduced by variations in weight of sample can be tremendous if very small or very large samples are used. Samples of 50 to 500 g of bone char can be satisfactorily sieved. However, it is recommended that samples of 100 to 150 g be used. The shaking time should be adjusted in accordance with the weight and particle size distribution of the sample.

The over-all reproducibility of a sieve analysis carried out in accordance with the following recommendations will be about 0.1 percent of the weight of the total sample. This variation will appear on all sieves irrespective of the amount of material retained on each individual sieve.

VI. Recommendations for Sieve Analyses of Bone Char

1. All testing sieves should be tested for uniformity of openings. 2. The effective openings of all testing sieves should be measured by means of a calibrated sample of material such as glass spheres, and the effective opening should be used instead of the nominal opening. All sieve analyses may then be calculated to correspond to the nominal openings. 3. If a Ro-Tap machine of the present design is used, it should be operated at 115 taps/min of the knocker mechanism. The other variables concerned with the shaking procedure are of minor importance and need not be rigidly controlled. 4. The sample should be dried before sieve analysis, although 1 or 2 percent of moisture in bone char has negligible effect. 5. The weight of the sample of bone char should be between 100 and 150 g, and the shaking time should be adjusted accordingly. 6. Fractions should be weighed to the nearest 0.1 g and recorded to the nearest 0.1 percent. For purposes of uniform comparison between laboratories it is recommended that cumulative percentage finer be used to express the results of sieve analyses.

VII. References

[1] Sture Mörtsell, Accuracy of sieve analyses made by means of sieving machines, (in English), Acta Polytech. No. 30 (1949).

[2] C. Fagerholt, Particle size distribution of products ground in tube mill, (in English) (GEC Gads, Copenhagen, Denmark, 1945).

[3] J. M. DalleValle, Micromeritics (Pitman Publishing Corp., New York, N. Y., 1943).

[4] Theodore Hatch, Determination of average particle size from the screen analysis of non-uniform particulate substances, J. Franklin Inst. **215**, 27 (1933).

[5] NBS Letter Circular LC584 (Mar. 1, 1940, amended May 10, 1948). See also, Federal Specification for Sieves; Standard, Testing R R–S–366a (Apr. 29, 1942). For sale by the Superintendent of Documents, Washington 25, D. C. Price 5 cents.

[6] Standard Specifications for sieves for testing purposes, ASTM Designation E 11–39, 1946 Book of ASTM Standards, Part II, p. 1237.

[7] Standard specifications for sieves for testing purposes, ASA No. Z23.1 (1939).

[8] M. Weber, Jr. and Raymond F. Moran, A precise method for sieve analyses, Ind. Eng. Chem. Anal. ed. **10**, 180 (1938).

[9] Standard method of test for sieve analysis of granular mineral surfacing for asphalt roofing and shingles, ASTM Designation D–451–40, 1946 Book of ASTM Standards, Part II, p. 740.

[10] Standard method of test for sieve analysis of non-granular mineral surfacing for asphalt roofing and shingles, ASTM Designation D 452–40, 1946 Book of ASTM Standards, part II, p. 743.

[11] Standard method of testing molding powders used in manufacturing molded electrical insulators, ASTM Designation D 392–38, 1946 Book of ASTM Standards, part III B, p. 6.

[12] Standard method of test for fineness of hydraulic cement by the No. 200 sieve, ASTM Designation C–184–44, 1946 Book of ASTM Standards, part II, page 59.

[13] Standard method of test for particle size of soaps and other detergents, ASTM Designation D 502–39, 1946 Book of ASTM Standards, part III A, p. 413.

[14] Standard method of sampling and fineness test of powdered coal, ASTM Designation D 197–30, 1946 Book of ASTM Standards, part III A, p. 50.

[15] Standard method of test for sieve analysis and water content of refractory materials, ASTM Designation C92–46, 1946 Book of ASTM Standards, part II, p. 325.

[16] Standard method of test for sieve analysis of mineral filler (road and paving materials), ASTM Designation D 546–41, 1946 Book of ASTM Standards, part II, p. 545.

[17] Standard method of test for sieve analysis of fine and coarse aggregates, ASTM Designation C–136–46, 1946 Book of ASTM Standards, part II, p. 467.

[18] George W. Snedecor, Statistical methods (Collegiate Press, Ames, Iowa, 1938).

[19] A. W. Fahrenwald and S. W. Stockdale, Effect of sieve motion on screening efficiency, U. S. Bur. Mines Rept. Investigations 2933 (1929).

VIII. Appendix

An analysis of variance for the top sieve takes the following form (for terminology, see any standard text on the analysis of variance, for example [18]):

Analysis of variance including all four knockers

[The value of F for significance at the 5-percent level is about 10 and at the 1-percent level about 30.]

Source of variation	Degrees of freedom	Sum of squares	Variance	F value
Knockers	3	3.275	1.085	93.0
Machines	1	0.0625	0.0625	5.36
Chars	1	131.1	131.1	11,230
Interactions: Machines and chars	1	0.16	0.16	13.7
Machines and knockers	3	2.0625	.6875	58.8
Chars and knockers	3	0.0525	.0175	1.5
Triple interaction (error)	3	.035	.01167	.....
Total	15	136.73	.....	.....

Since it is noted that the absence of a knocker (condition knocker 4) is the greatest cause of the variation, the above analysis of variance is repeated omitting K_4:

Analysis of variance omitting K_4 (no knocker)

[The value of F for significance at the 5-percent level is about 20 and at the 1-percent level about 100.]

Source of variation	Top sieve		Second sieve		Third sieve		Pan	
	Variance	F	Variance	F	Variance	F	Variance	F
Knockers	0.0258	1.49	0.01	3	0.2580	1.2	0.01	.....
Machines	.8533	49	.0208	6.2	.1633	5.3	.48	.....
Chars	98.06	5,600	6.02	1,800	6.45	209	.03	.....
Interactions: Machines and chars	0.120	6.8	0.0208	6.2	0.0533	1.8	.12	.....
Machines and knockers	.0308	1.8	.0133	4	.0108	2.8	.01	.....
Chars and knockers	.0583	1.5	.0133	4	.0358	1.2	0	.....
Triple interaction (error)	.0175	.....	.0033	.....	.0308	.....	0	.....

WASHINGTON, March 21, 1950.

Journal of Research of the National Bureau of Standards Vol. 45, No. 4, October 1950 Research Paper 2144

First Spectrum of Arsenic

By William F. Meggers, Allen G. Shenstone,[1] and Charlotte E. Moore

The spectrum emitted by neutral arsenic atoms was observed photographically in the infrared, visible, and ultraviolet, and new lines were discovered in each spectral region. Measured wavelengths and estimated relative intensities are given for 330 lines, ranging from 1407.34 to 11679.9 A in wavelength and from 1 to 2000 in intensity. More than 74 percent of the total number and 97 percent of the total intensity of observed lines have been explained as combinations of 30 odd energy levels arising from $4s^2\,4p^3$ and $4s^2\,4p^2\,np$ electron configurations and 58 even levels from $4s\,4p^4$, $4s^2\,4p^2\,ns$, and $4s^2\,4p^2\,nd$. The average difference between observed and computed wave numbers is 0.14 cm^{-1}. Most of the observed levels have been assigned to doublet and quartet terms, and spectral series of the type $4s^2\,4p^3 - 4s^2\,4p^2\,ns$ have been identified. Calculations based on these series yield an absolute value of 79165 cm^{-1} for the ground state $4s^2\,4p^3\,{}^4S^\circ_{3/2}$ of neutral arsenic atoms, that is, an ionization potential of 9.81 $\pm$0.01 electron volts.

I. Introduction

In 1929 Meggers and deBruin [1][2] published a paper on the arc spectrum of arsenic, based on measurements of 54 ultraviolet lines (1889.85 to 3119.60 A) and 23 infrared lines (7410.07 to 10023.98 A). At that time this spectrum had not been observed in the region of shorter waves, and the assumption that the Blochs [2] had observed As I lines in spark spectra in the extreme ultraviolet could not be verified. Although it was not possible then to determine absolute term values from spectral series, the ground state was recognized as $^4S^\circ$, and by comparison with analogous terms in the spectra of neighboring elements the absolute value of this ground state was tentatively estimated as 80693 cm^{-1}, which corresponded to an ionization potential of about 10 ev. This was 13 percent lower than the value 11.54 $\pm$0.5 ev derived in 1922 by Ruark, et al. [3] from experiments on low-voltage arcs in arsenic vapor.

A paper on spectra of arsenic in the extreme ultraviolet, 2500 to 710 A, by Queney [4] reported about 300 arsenic lines from electrodeless discharges but added nothing to the As I spectrum.

The shorter waves of this spectrum were first investigated by K. R. Rao [5], who burned metallic arsenic in an arc between carbon poles and photographed the spectrum with an evacuated spectrograph containing a 1 m-radius grating, giving a scale of 8.6 A/mm. He measured 64 lines between 1995.45 and 1563.08 A with estimated probable errors of $\pm$0.03 A. Because no spectral series could be found, Rao adopted the absolute value $^4S^\circ = 93500$ cm^{-1} from the ionization potential 11.54 ev reported by Ruark, et al. [3].

In 1932 further investigations of the arc spectrum of arsenic were reported by A. S. Rao [6], who measured 160 ultraviolet As I lines (1995.45 to 1319.48 A) on a hollow-cathode spectrogram having a scale of

[1] Department of Physics, Princeton University, Princeton, N. J.
[2] Figures in brackets indicate the literature references at the end of this paper.

UNIVERSITY OF MICHIGAN LIBRARIES

17 A/mm. He quoted infrared data from Meggers and deBruin [1], ultraviolet data (3119.576 to 1563.08 A) from K.R. Rao [5], and classified 200 lines with 66 levels. Earlier interpretations of the As I spectrum were revised and extended, and an estimate of the ionization potential as 10.5 ev was derived from two members of several assumed series terms. In the light of present data, many of Rao's designations of spectral terms appear to be incorrect, the reality of some of his levels is doubtful, and his value of the ionization potential for arsenic is certainly too large.

In 1933 Hicks [7] attempted to obtain an absolute value of the spectral terms of arsenic by "the application of the oun-multiple law to the separations". Mysterious calculations led him to a value of 87945 cm^{-1} for the ionization limit, and he stated that "this is equivalent to 10.85 volts, and closer to the value 11.52 $\pm$.5 referred to by Rao than his 85000 with 10.5 volts". It now appears that this correction was wrong in both direction and magnitude.

The only other investigation of the first spectrum of arsenic published during the past 20 years deals with the Zeeman effect. Green and Barrows [8], in 1935, published Zeeman patterns for 11 lines (2288.2 to 3033.0 A), derived *g*-values for 13 levels, and concluded that "the levels of As I are so close to *LS* coupling that the *g*-values are almost normal".

The lack of a reliable ionization potential was inspiration and justification for another investigation of the first spectrum of arsenic. Such an investigation has been made at the National Bureau of Standards from time to time during the past 20 years, and as a result, the As I spectrum has been exhaustively explored photographically from 12500 A in the infrared to 1200 A in the ultraviolet. This spectrum has been found to consist of three separated groups of lines, 71 infrared, 46 visible, and 210 ultraviolet. More than 74 percent of the total number and 97 percent of the total intensity of observed As I lines have been explained as combinations of 88 atomic-energy levels, and a principal ionization potential of 9.81 ev has been calculated with confidence. Because further investigation of this spectrum is neither likely nor profitable, this final report is now presented.

II. Wavelength Measurements

The infrared As I lines reported by Meggers and deBruin [1] were observed with photographic plates sensitized in the laboratory by bathing them in dilute solutions of neocyanine, and the greatest wavelength that could be recorded was 10024 A. In 1934 "Xenocyanine Plates" prepared by the Eastman Kodak Co. permitted an extension of the infrared recording of As I lines to 10888 A, and in 1940 Eastman I-Z plates were used to explore this spectrum to 12500 A. No As I lines were recorded beyond 11680 A, but this wavelength is more than 1000 A greater than the limit observed radiometrically by Randall [9] in 1911.

The light sources, spectrographs, and standard wavelengths employed in observing the infrared, visible, and air-transparent ultraviolet As I spectrum were described in detail by Meggers and deBruin [1].

A number of visible lines not previously noted in the arc spectrum of arsenic and not identifiable with impurities were observed in 1929 [1], but few details were given because no structural connection was found between the visible lines and the invisible ones. Such a connection was first found in 1932 by deBruin,[3] who saw that some of the visible lines exhibited the same wave-number differences as some infrared ones. At that time it was not possible to classify and interpret all the visible lines, but this was facilitated by the more recent extensions of the infrared data.

The most important improvements of As I data were made in the extreme ultraviolet, from arc spectrograms exposed at Princeton University and measured at the National Bureau of Standards. The spectrograms were made with a normal incidence vacuum spectrograph having a 2 m-radius glass grating ruled 30,000 lines per inch. The dispersion is 4.2 A/mm and the definition and resolving power are excellent, but Rowland ghosts are prominent on strong exposures. All the identified ghosts contributed to determinations of the wavelengths of their respective parent lines.

Our first attempt to observe the short waves of the As I spectrum was made in 1940 with electrodes of 25 percent As and 75 percent Cu prepared in the Metallurgy Division of this Bureau. Four spectrograms were measured relative to copper standards, but only 100 As I lines were found between 1995.43 and 1472.38 A. Because the arsenic spectrum could not be fully developed with this alloy without greatly overexposing the copper spectra, additional spectrograms were made in 1949 with metallic arsenic in a cupped carbon electrode opposite a copper electrode. These were highly satisfactory for the final description of the As I spectrum to its ultraviolet limit near 1400 A; after eliminating lines due to As II, impurities, and ghosts, they yielded 168 As I lines between 1995.43 and 1407.34 A. These wavelengths were measured relative to Cu I and Cu II lines given in Shenstone's papers [10] supplemented by carbon, nitrogen, and oxygen "impurity" lines recommended as tentative standards by Boyce and Robinson [11]. All except the weak lines were measured on six or seven spectrograms, and the agreement between determinations, as well as the agreement between values observed and calculated from atomic-energy levels, prove that the probable error rarely exceeds $\pm$0.01 A. The complete list of observed lines was compared with the Table of Principal Lines of all Elements [12] for the purpose of identifying possible impurities, and the ultraviolet portion was also compared with a description by Rao [13] of the first spark spectrum of arsenic for the purpose of eliminating about 10 As II lines that appeared in the arc spectrum.

Measured wavelengths and estimated intensities of 330 lines characteristic of neutral arsenic atoms are

[3] Communicated by T. L. deBruin to W. F. Meggers in a note dated Feb. 27, 1932.

presented in the first and second columns, respectively, of table 1. The wavelengths greater than 2000 A are valid for standard air; those smaller than 2000 A are vacuum values. The intensity numbers are assigned from visual comparisons of the blackness and width of slit images on the most strongly exposed spectrograms; they are comparable only for lines in limited regions of spectra and obviously decline rapidly beyond 11000A as photographic sensitivity decreases for increasing wavelength. In addition to intensity, the character of certain lines is indicated by symbols: R=wide self-reversal, r=narrow self-reversal, h=hazy, e=observed at electrode.

TABLE 1. *First spectrum of arsenic* (As I)

λ air A	Intensity, character	Wave No. vac. cm⁻¹	Obs.−Calc. Wave No. 0.1 cm⁻¹	Term combination	λ air A	Intensity, character	Wave No. vac. cm⁻¹	Obs.−Calc. Wave No. 0.1 cm⁻¹	Term combination
1	2	3	4	5	1	2	3	4	5
11679.9	3	8559.37			7863.43	4	12713.61	0	$5s\ ^4P_{1/2}-5p\ ^2P^0_{3/2}$
11520.9	10	8677.54	0	$5s\ ^2P_{1/2}-5p\ ^4P^0_{3/2}$	7829.82	2	12768.17	+1	$5p\ ^4D^0_{3/2}-4d''\ ^2D_{3/2}$
11377.1	4	8787.19	+1	$5s\ ^4P_{3/2}-5p\ ^4D^0_{5/2}$	7648.15	2	13071.46		
11244.5	10	8890.79	+1	$5s\ ^2P_{3/2}-5p\ ^4P^0_{5/2}$	7410.02	8	13491.53	+1	$5p\ ^4D^0_{3/2}-75176$
10888.82	20	9181.22	−1	$5s\ ^4P_{1/2}-5p\ ^4D^0_{3/2}$	7384.97	10	13537.29		
10808.11	50	9249.78	0	$5s\ ^4P_{1/2}-5p\ ^4D^0_{1/2}$	7384.62	10	13537.94		
10793.37	8	9262.42	0	$5s\ ^2P_{3/2}-5p\ ^4P^0_{3/2}$	6338.94	40	15771.16	−1	$5s\ ^2P_{1/2}-5p'\ ^2D^0_{3/2}$
10780.36	5	9273.59	0	$5s'\ ^2D_{3/2}-5p'\ ^2D^0_{3/2}$?	6176.68	1	16185.46		
10614.07	200	9418.87	+1	$5s\ ^2P_{3/2}-5p\ ^2D^0_{3/2}$	6049.48	5h	16525.78	+1	$5s\ ^2P_{1/2}-6p\ ^4D^0_{3/2}$
10575.02	100	9453.65	0	$5s\ ^2P_{1/2}-5p\ ^2S^0_{1/2}$	5815.40	2h	17190.97	+1	$5s\ ^2P_{1/2}-5p'\ ^2P^0_{1/2}$
10490.29	2h	9530.02			5751.85	5h	17380.90	−1	$5s\ ^2P_{3/2}-6p\ ^4D^0_{1/2}$
10477.30	4	9541.83	0	$5s'\ ^2D_{1/2}-5p'^2D^0_{3/2}$	5739.54	2h	17418.18		
10455.64	20	9561.60	0	$5s'\ ^2D_{3/2}-5p'^2D^0_{3/2}$	5534.47	3	18063.57	+1	$5s\ ^2P_{3/2}-6p\ ^2S^0_{1/2}$
10453.09	100	9563.93	0	$5s\ ^2P_{1/2}-5p\ ^2D^0_{3/2}$	5497.10	4	18186.37		
10445.6	3h	9570.84	+2	$5p\ ^2D^0_{3/2}-4d''\ ^2P_{1/2}$	5494.40	5	18195.31	0	$5s\ ^2P_{1/2}-6p\ ^2P^0_{3/2}$
10325.74	10	9681.88	+1	$5s'\ ^2D_{1/2}-6p\ ^4D^0_{1/2}$	5451.32	150	18339.10	0	$5s\ ^2P_{1/2}-6p\ ^2D^0_{3/2}$
10304.82	5	9701.54	−1	$5s'\ ^2D_{3/2}-6p\ ^4D^0_{1/2}$	5422.66	2	18436.02		
10286.85	50	9718.49	0	$5s\ ^2P_{1/2}-5p\ ^2P^0_{3/2}$	5408.13	100	18485.56	−2	$5s\ ^2P_{3/2}-6p\ ^2D^0_{3/2}$
10024.04	400	9973.30	0	$5s\ ^4P_{3/2}-5p\ ^4D^0_{5/2}$	5363.54	60	18639.23	−3	$5s\ ^2P_{1/2}-6p\ ^2P^0_{1/2}$
10010.63	100	9986.65	0	$5s'\ ^2D_{1/2}-5p'\ ^2F^0_{3/2}$	5361.12	10	18647.65		
9990.87	5	10006.40	0	$5s'\ ^2D_{3/2}-5p'\ ^2F^0_{3/2}$	5357.42	1h	18660.53	−1	$5s\ ^2P_{3/2}-5p'\ ^2P^0_{3/2}$
9923.05	400	10074.78	0	$5s\ ^4P_{1/2}-5p\ ^4D^0_{3/2}$	5318.76	5	18796.16	+1	$5s\ ^2P_{3/2}-5p'\ ^2P^0_{1/2}$
9915.71	200	10082.24	−1	$5s\ ^2P_{1/2}-5p\ ^2P^0_{1/2}$	5212.14	5	19180.65		
9900.55	150	10097.68	0	$5s\ ^4P_{3/2}-5p\ ^4D^0_{3/2}$	5210.23	20	19187.68	0	$5s\ ^4P_{3/2}-6p\ ^4P^0_{3/2}$
					5203.84	4	19211.24	0	$5s\ ^4P_{1/2}-5p'\ ^4F^0_{3/2}$
9886.05	100	10112.49	0	$5s'\ ^2D_{3/2}-5p'\ ^2F^0_{1/2}$					
9869.5	2	10129.5			5190.20	50	19259.49	0	$5s\ ^4P_{1/2}-6p\ ^4P^0_{1/2}$
9850.4	1	10149.1			5141.63	100	19443.68	0	$5s\ ^4P_{5/2}-6p\ ^4D^0_{3/2}$
9833.76	300	10166.26	+1	$5s\ ^4P_{3/2}-5p\ ^4D^0_{1/2}$	5130.78	80	19484.80	+1	$5s\ ^2P_{3/2}-6p\ ^4S^0_{1/2}$
9781.32	30	10220.77			5121.34	100	19520.72	−1	$5s\ ^4P_{1/2}-6p\ ^4D^0_{1/2}$
					5105.55	30 Cu?	19581.09	0	$5s\ ^4P_{5/2}-6p\ ^4D^0_{5/2}$
9772.08	7	10230.43			5103.53	20	19588.84	−1	$5s\ ^4P_{1/2}-6p\ ^2S^0_{1/2}$
9721.8	1	10283.3							
9714.64	10	10290.92			5099.59	60	19603.97	0	$5s\ ^4P_{3/2}-6p\ ^4P^0_{3/2}$
9690.83	10	10316.20	+2	$5s'\ ^2D_{5/2}-6p\ ^4D^0_{5/2}$	5083.76	10	19665.02	0	$5s\ ^2P_{3/2}-6p\ ^2P^0_{3/2}$
9654.54	10	10354.98			5068.98	100	19722.35	0	$5s\ ^4P_{3/2}-6p\ ^4S^0_{3/2}$
					5043.31	50	19822.74	−1	$5s\ ^4P_{5/2}-6p\ ^4D^0_{3/2}$
9648.06	20	10361.94			4995.83	8	20011.13	−1	$5s\ ^4P_{1/2}-6p\ ^2D^0_{3/2}$
9645.8	3	10364.4	+1	$5s'\ ^2D_{1/2}-6p\ ^2S^0_{1/2}$					
9626.70	400	10384.93	0	$5s\ ^4P_{3/2}-5p\ ^4P^0_{3/2}$	4987.02	6A	20046.48	0	$5s\ ^4P_{3/2}-6p\ ^2D^0_{5/2}$
9597.95	300	10416.04	−1	$5s\ ^4P_{1/2}-5p\ ^4P^0_{1/2}$	4952.56	10	20185.96	0	$5s\ ^4P_{1/2}-5p'\ ^2P^0_{1/2}$
9510.9	3	10511.4	0	$5s\ ^2P_{3/2}-5p\ ^4S^0_{3/2}$	4919.49	2	20321.65	+1	$5s\ ^4P_{1/2}-5p'\ ^2P^0_{3/2}$
					4910.23	10h	20359.98		
9300.61	250	10749.04	−1	$5s\ ^4P_{5/2}-5p\ ^4S^0_{3/2}$	4873.78	2h	20512.24		
9287.28	150	10787.70	−1	$5s\ ^4P_{1/2}-5p\ ^4P^0_{1/2}$	4785.27	3h	20891.64	−1	$5s\ ^4P_{1/2}-6p\ ^4P^0_{1/2}$
			+11	$(5s'\ ^2D_{1/2}-6p\ ^2D^0_{3/2})$	4777.02	1h	20927.72	+1	$5s\ ^4P_{5/2}-6p\ ^2D^0_{3/2}$
9165.6	3 Cu?	10907.4			4758.28	5h	21010.14	0	$5s\ ^4P_{1/2}-6p\ ^4S^0_{1/2}$
9134.78	50	10944.17	0	$5s\ ^4P_{1/2}-5p\ ^2D^0_{3/2}$	4622.82	6A	21625.78		
9120.32	10	10961.52	+1	$5s'\ ^2D_{3/2}-5p'\ ^2P^0_{3/2}$	4621.27	4h	21633.03		
9027.8	2	11073.9							
9008.95	2	11097.02	+1	$5s'\ ^2D_{3/2}-5p'\ ^2P^0_{1/2}$	4355.39	10h	22953.63		
8993.05	20	11116.65	−1	$5s'\ ^2D_{3/2}-5p'\ ^2P^0_{3/2}$	4342.05	10h	23024.15		
8980.13	5	11132.64	−3	$5p\ ^4D^0_{3/2}-4d''\ ^2P_{3/2}$	4336.15	5h	23055.47		
8935.56	50	11188.17	0	$5s\ ^2P_{3/2}-5p\ ^2P^0_{3/2}$	4320.12	3h	23141.02		
					4313.16	4h	23178.36		
8874.8	2h	11264.8	−2	$5p\ ^4D^0_{1/2}-4d''\ ^2P_{1/2}$					
8869.66	100	11271.30	0	$5s\ ^4P_{5/2}-5p\ ^4D^0_{5/2}$	4301.56	5hl	23240.87		
8821.73	150	11332.41	−1	$5s\ ^4P_{3/2}-5p\ ^4P^0_{3/2}$	3119.60	50	32046.1	−2	$4p^3\ ^2P^0_{1/2}-5s\ ^4P_{3/2}$
8654.14	100	11551.99	0	$5s\ ^2P_{3/2}-5p\ ^2P^0_{1/2}$	3075.32	20	32507.5	−2	$4p^3\ ^2P^0_{3/2}-5s\ ^4P_{5/2}$
8646.03	3	11562.83			3032.85	40	32962.7	0	$4p^3\ ^2P^0_{3/2}-5s\ ^4P_{1/2}$
					2990.99	20	33424.0	−1	$4p^3\ ^2P^0_{5/2}-5s\ ^4P_{3/2}$
8564.71	100	11672.62	0	$5s\ ^4P_{1/2}-5p\ ^4P^0_{1/2}$	2918.82	2	34250.4	0	$4p^3\ ^2P^0_{1/2}-5s\ ^4P_{5/2}$
8541.60	50	11704.20	0	$5s\ ^4P_{3/2}-5p\ ^4P^0_{1/2}$	2898.71	50r	34488.0	−1	$4p^3\ ^2P^0_{1/2}-5s\ ^2P^0_{3/2}$
8455.1	8	11823.9			2860.44	100r	34949.4	−1	$4p^3\ ^2P^0_{3/2}-5s\ ^2P^0_{3/2}$
8428.91	100	11860.66	+1	$5s\ ^4P_{5/2}-5p\ ^2D^0_{3/2}$	2780.22	200r	35957.8	0	$4p^3\ ^2P^0_{3/2}-5s\ ^2P^0_{1/2}$
8374.84	10	11935.82	0	$5s'\ ^2D_{3/2}-6p\ ^2P^0_{3/2}$	2745.00	50r	36419.1	−1	$4p^3\ ^2P^0_{5/2}-5s\ ^2P^0_{1/2}$
8305.61	50	12036.75	0	$5s\ ^4P_{1/2}-5p\ ^4S^0_{3/2}$	2492.91	50	40101.7	+4	$4p^3\ ^2D^0_{1/2}-5s\ ^4P_{5/2}$
8242.15	20	12129.43	0	$5s'\ ^2D_{5/2}-6p\ ^2D^0_{5/2}$	2456.53	200r	40695.5	−1	$4p^3\ ^2D^0_{3/2}\ \ 5s\ ^4P_{1/2}$
8055.72	5h	12410.13	+1	$5s'\ ^2D_{1/2}-6p\ ^2P^0_{1/2}$	2437.23	50	41017.7	0	$4p^3\ ^2D^0_{3/2}-5s\ ^4P_{1/2}$
8042.95	8	12429.83	0	$5s'\ ^2D_{3/2}-6p\ ^2P^0_{1/2}$	2381.18	150r	41983.2	−1	$4p^3\ ^2D^0_{3/2}-5s\ ^4P_{3/2}$
7960.27	30	12558.94	−1	$5s\ ^4P_{1/2}-5p\ ^2D^0_{3/2}$	2370.77	100r	42167.5	0	$4p^3\ ^2P^0_{1/2}-5s'\ ^2D_{3/2}$

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 1. *First spectrum of arsenic* (As I)—Continued

λ_{air} Å	Intensity, character	Wave No. vac. cm^{-1}	Obs.−Calc. Wave No. 1.0 cm^{-1}	Term combination	$\lambda_{vac.}$ Å	Intensity, character	Wave No. vac. cm^{-1}	Obs.−Calc. Wave No. 0.1 cm^{-1}	Term combination
1	2	3	4	5	1	2	3	4	5
2369.67	80r	42187.1	−2	$4p^3\ ^2P_{1/2}-5s'\ ^2D_{3/2}$	1844.57	40	54213.2	−1	$4p^3\ ^2P_{3/2}-5s''\ ^2S_{1/2}$
2363.05	10	42305.2	−2	$4p^3\ ^2D_{1/2}-5s\ ^4P_{1/2}$	1844.36	40	54219.3	+3	$4p^3\ ^2D_{1/2}-4p^4\ ^2P_{1/2}$
2349.84	500R	42543.0	−1	$4p^3\ ^2D_{1/2}-5s\ ^2P_{3/2}$	1840.48	10	54333.6	+3	$4p^3\ ^2P_{3/2}-72519$
2344.03	50	42648.5	−2	$4p^3\ ^2P_{3/2}-5s'\ ^2D_{3/2}$	1831.74	30	54592.9	0	$4p^3\ ^2P_{3/2}-72779$
2288.12	500R	43690.5	−2	$4p^3\ ^2D_{3/2}-5s\ ^2P_{1/2}$	1831.30	50	54606.0	+7	$4p^3\ ^4S_{3/2}-5s\ ^2P_{1/2}$
2271.36	50	44012.8	0	$4p^3\ ^2D_{1/2}-5s\ ^2P_{1/2}$	1821.32	2	54905.2	+3	$4p^3\ ^2D_{1/2}-4p^4\ ^2P_{3/2}$
2266.70	25	44103.3	−3	$4p^3\ ^2P_{1/2}-4d\ ^4P_{3/2}$	1818.59	5h	54987.7	+2	$4p^3\ ^2P_{1/2}-73635$
2228.66	20	44856.0	−1	$4p^3\ ^2P_{1/2}-4d\ ^4P_{1/2}$	1816.72	10	55044.3		
2205.97	15	45317.4	−1	$4p^3\ ^2P_{3/2}-4d\ ^4P_{1/2}$	1806.15	200	55366.4	0	$4p^3\ ^4S_{3/2}-4p^4\ ^4P_{3/2}$
2205.16	10	45334.0	−2	$4p^3\ ^2P_{1/2}-4d\ ^4P_{3/2}$					
					1803.46	10h	55449.0	+1	$4p^3\ ^2P_{3/2}-73635$
2198.34	5	45474.6	+1	$4p^3\ ^2P_{1/2}-4d\ ^2P_{3/2}$	1801.92	3	55496.4	+1	$4p^3\ ^2P_{1/2}-74143$
2187.75	5	45694.7	−2	$4p^3\ ^2P_{1/2}-4d\ ^2D_{3/2}$	1799.51	2	55570.7	+1	$4p^3\ ^2D_{3/2}-4d\ ^2P_{1/2}$
2182.94	20	45795.5	−1	$4p^3\ ^2P_{3/2}-4d\ ^4P_{3/2}$	1798.61	9	55598.5		
2176.26	5	45935.9	0	$4p^3\ ^2P_{3/2}-4d\ ^2P_{3/2}$	1791.77	40	55810.7	−1	$4p^3\ ^2D_{3/2}-4d\ ^2D_{3/2}$
2165.52	150	46163.8	−2	$4p^3\ ^2P_{1/2}-4p^4\ ^2P_{1/2}$					
					1789.85	50	55870.6	0	$4p^3\ ^2D_{3/2}-4d\ ^2F_{1/2}$
2144.08	100	46625.3	−1	$4p^3\ ^2P_{3/2}-4p^4\ ^2P_{1/2}$	1789.14	15	55892.8	+1	$4p^3\ ^2D_{1/2}-4d\ ^2P_{1/2}$
2142.80	2	46653.2			1787.07	6	55957.5	−2	$4p^3\ ^2P_{3/2}-74143$
2133.80	50	46849.9	0	$4p^3\ ^2P_{1/2}-4p^4\ ^2P_{3/2}$	1785.84	3	55996.1	+3	$4p^3\ ^2D_{1/2}-4p^4\ ^2S_{1/2}$
2112.99	100	47311.3	0	$4p^3\ ^2P_{3/2}-4p^4\ ^2P_{3/2}$	1781.48	50	56133.1	+2	$4p^3\ ^2D_{1/2}-4d\ ^2D_{3/2}$
2089.74	6	47837.6	−1	$4p^3\ ^2P_{3/2}-4d\ ^2P_{1/2}$					
					1780.52	50	56163.4	+4	$4p^3\ ^2P_{1/2}-4d''\ ^2D_{3/2}$
2085.25	30	47940.6	−2	$4p^3\ ^2P_{1/2}-4p^4\ ^2S_{1/2}$	1775.12	2	56334.2		
2079.30	7	48077.7	−2	$4p^3\ ^2P_{1/2}-4d\ ^2D_{3/2}$	1772.54	8	56416.2	0	$4p^3\ ^2D_{3/2}-6s\ ^4P_{1/2}$
2069.78	30	48298.8	−3	$4p^3\ ^2P_{3/2}-4d\ ^2P_{1/2}$	1771.84	2	56438.5	0	$4p^3\ ^2P_{1/2}-75086$
2067.11	20	48361.2	0	$4p^3\ ^2P_{1/2}-6s\ ^4P_{3/2}$	1768.97	5	56530.1	+12	$4p^3\ ^2P_{1/2}-75176?$
2065.36	50	48402.2	0	$4p^3\ ^2P_{3/2}-4p^4\ ^2S_{1/2}$					
					1767.97	6e,h?	56562.0		
2061.61	3	48490.2			1758.60	100	56863.6	0	$4p^3\ ^4S_{1/2}-4p^4\ ^4P_{1/2}$
2047.57	50	48822.6	0	$4p^3\ ^2P_{3/2}-6s\ ^4P_{3/2}$	1757.47	4	56900.0	+1	$4p^3\ ^2P_{3/2}-75086$
2028.86	2	49272.8	+1	$4p^3\ ^2P_{1/2}-6s\ ^4P_{1/2}$	1756.51	2h	56931.1	−1	$4p^3\ ^2P_{1/2}-75578$
2024.34	5	49382.8			1754.21	10	57005.7	+1	$4p^3\ ^2D_{3/2}-6s\ ^4P_{1/2}$
2013.32	100	49653.1	0	$4p^3\ ^2P_{1/2}-4p^4\ ^2D_{1/2}$					
					1750.37	2	57130.8		
2012.76	15	49666.9	−1	$4p^3\ ^2P_{1/2}-6s\ ^2P_{3/2}$	1749.72	1e,h	57152.0	+1	$4p^3\ ^2P_{1/2}-75799$
2010.04	20	49734.0	−1	$4p^3\ ^2P_{3/2}-6s\ ^4P_{1/2}$	1745.60	1	57286.9		
2009.19	100r	49755.1	−1	$4p^3\ ^2P_{1/2}-4p^4\ ^2D_{3/2}$	1742.59	10	57385.8	−2	$4p^3\ ^2D_{3/2}-4p^4\ ^2D_{1/2}$
2003.34	300r	49900.4	0	$4p^3\ ^2D_{3/2}-5s'\ ^2D_{3/2}$	1741.28	10	57429.0	+3	$4p^3\ ^2P_{1/2}-76076$
2002.54	20	49920.3	+1	$4p^3\ ^2D_{3/2}-5s'\ ^2D_{1/2}$					
vac.					1740.93	3	57440.6		
1995.43	100r	50114.5	0	$4p^3\ ^2P_{3/2}-4p^4\ ^2D_{3/2}$	1739.49	60	57488.1	0	$4p^3\ ^2D_{3/2}-4p^4\ ^2D_{3/2}$
1994.89	20	50128.3	−1	$4p^3\ ^2P_{3/2}-6s\ ^2P_{3/2}$				0	$4p^3\ ^4S_{3/2}-4p^4\ ^4P_{3/2}$
1991.13	100r	50222.7	+2	$4p^3\ ^2D_{1/2}-5s'\ ^2D_{3/2}$	1735.70	3	57613.6	+3	$4p^3\ ^2P_{3/2}-75799$
1990.35	200r	50242.4	+1	$4p^3\ ^2D_{1/2}-5s'\ ^2D_{1/2}$	1732.86	30	57708.1	0	$4p^3\ ^2D_{3/2}-4p^4\ ^2D_{3/2}$
1973.65	3	50667.5	+2	$4p^3\ ^2P_{1/2}-6s\ ^2P_{1/2}$	1732.44	30	57722.1	+1	$4p^3\ ^2D_{3/2}-6s\ ^2P_{3/2}$
1972.62	1000R	50694.0	+2	$4p^3\ ^4S_{3/2}-5s\ ^4P_{3/2}$	1729.80	30	57810.2	0	$4p^3\ ^2D_{1/2}-4p^4\ ^2D_{3/2}$
1960.89	2	50997.2	−1	$4p^3\ ^2P_{1/2}-69644$	1727.38	4	57891.1	+10	$4p^3\ ^2P_{3/2}-76076?$
1958.91	40r	51048.8	+1	$4p^3\ ^2P_{3/2}-6s\ ^2P_{1/2}$	1724.77	7	57978.7		
1958.82	20	51051.1	+3	$4p^3\ ^2P_{1/2}-69698$	1722.37	4e	58059.5		
1950.38	2	51272.1			1718.55	1	58188.6	+5	$4p^3\ ^2P_{1/2}-76835$
1943.30	2e	51458.9	+2	$4p^3\ ^2P_{3/2}-69644$	1712.32	10	58400.3	+1	$4p^3\ ^2D_{3/2}-6s\ ^4P_{3/2}$
1941.36	5	51510.3	+2	$4p^3\ ^2P_{3/2}-6s\ ^2P_{1/2}$	1710.16	7h	58474.1	−1	$4p^3\ ^2P_{1/2}-77121$
1937.59	1500R	51610.4	+2	$4p^3\ ^4S_{3/2}-5s\ ^4P_{1/2}$	1709.26	2h	58504.8		
1929.14	3	51836.6	+1	$4p^3\ ^2D_{3/2}-4d\ ^4P_{3/2}$	1709.03	2	58512.7		
1928.30	1	51859.2			1705.74	4	58625.6	+6	$4p^3\ ^2P_{1/2}-77272$
1919.72	1	52090.9			1702.94	3	58722.0	−3	$4p^3\ ^2D_{1/2}-6s\ ^4P_{1/2}$
1918.39	3	52127.0			1701.22	30	58781.3	−3	$4p^3\ ^2D_{3/2}-6s\ ^2P_{1/2}$
1917.21	20	52159.1	+5	$4p^3\ ^2D_{1/2}-4d\ ^4P_{3/2}$	1701.16	15	58783.4	−3	$4p^3\ ^2D_{3/2}-69698$
1908.13	1	52407.3	−4	$4p^3\ ^2P_{1/2}-71055$	1691.87	7	59106.2	+4	$4p^3\ ^2D_{1/2}-69698$
1902.31	6	52567.7			1687.12	3e	59272.6		
1901.54	5	52589.0	0	$4p^3\ ^2D_{3/2}-4d\ ^4P_{1/2}$	1684.63	2e	59360.2		
1891.47	5	52868.9	−2	$4p^3\ ^2P_{3/2}-71055$	1681.47	1	59471.8		
1890.42	2000R	52898.2	+3	$4p^3\ ^4S_{3/2}-5s\ ^4P_{3/2}$	1679.14	2	59554.3		
1889.95	5	52911.4	+3	$4p^3\ ^2D_{1/2}-4d\ ^4P_{1/2}$	1674.59	3	59716.1		
1881.96	40	53136.1	+5	$4p^3\ ^4S_{3/2}-5s\ ^2P_{3/2}$	1671.88	2	59812.9		
1876.98	2	53277.1	+2	$4p^3\ ^2P_{1/2}-4d'\ ^2P_{3/2}$	1670.79	4	59851.9		
1873.02	40	53389.6	+4	$4p^3\ ^2D_{1/2}-4d\ ^4P_{3/2}$	1670.12	2	59875.9		
1871.68	30	53427.9	+1	$4p^3\ ^2D_{3/2}-4d\ ^2D_{3/2}$	1662.76	2	60141.0	+4	$4p^3\ ^2D_{3/2}-71055$
1871.14	2	53443.4			1653.92	4	60462.4	−3	$4p^3\ ^2D_{1/2}-71055$
1869.94	15	53477.6	+1	$4p^3\ ^2P_{1/2}-4d'\ ^2P_{1/2}$	1649.55	2	60622.6		
1865.10	8	53616.6			1644.33	5	60815.0	0	$4p^3\ ^4S_{3/2}-5s'\ ^2D_{3/2}$
1860.87	5	53738.3	0	$4p^3\ ^2P_{3/2}-4d'\ ^2P_{3/2}$	1643.79	2	60835.0	+2	$4p^3\ ^4S_{3/2}-5s'\ ^2D_{1/2}$
1860.46	80	53750.1	+2	$4p^3\ ^2D_{3/2}-4d\ ^2D_{1/2}$	1633.71	5	61210.4	0	$4p^3\ ^2D_{3/2}-4d'\ ^2P_{3/2}$
1860.40	80	53751.9	0	$4p^3\ ^2P_{1/2}-5s''\ ^2S_{1/2}$	1630.48	4	61331.6	−3	$4p^3\ ^2D_{1/2}-4d'\ ^2P_{3/2}$
1856.24	2	53872.3	+4	$4p^3\ ^2P_{3/2}-72519$	1625.17	1	61532.0	−5	$4p^3\ ^2D_{1/2}-4d'\ ^2P_{1/2}$
1855.39	10	53897.0	+1	$4p^3\ ^2D_{3/2}-4p^4\ ^2P_{1/2}$	1623.26	2	61604.4	−4	$4p^3\ ^2D_{3/2}-72519$
1853.95	3	53938.9	0	$4p^3\ ^2P_{3/2}-4d'\ ^2P_{1/2}$	1620.94	5	61692.6		
1853.21	20	53960.4			1616.44	6	61864.5	+1	$4p^3\ ^2D_{3/2}-72779$
1850.24	40	54047.0	10	$4p^3\ ^2D_{3/2}-4d\ ^2F_{3/2}$	1614.82	5	61926.5	−4	$4p^3\ ^2D_{1/2}-72519$
1847.32	10	54132.5	+0	$4p^3\ ^2P_{1/2}-72779$	1612.57	4	62012.8		

TABLE 1. *First spectrum of arsenic* (As I)—Continued

λ_{vac} A	Intensity, character	Wave No. vac. cm^{-1}	Obs.−Calc. Wave No. 0.1 cm^{-1}	Term combination	λ_{vac} A	Intensity, character	Wave No. vac. cm^{-1}	Obs.−Calc. Wave No. 0.1 cm^{-1}	Term combination
1	2	3	4	5	1	2	3	4	5
1608.07	3	62186.3	−2	$4p^3\ ^2D^o_{5/2}-72779$	1516.97	1*r*	65920.9	−1	$4p^3\ ^2D^o_{3/2}-76835$
1597.48	2	62598.6			1515.48	20*h*	65985.7		
1596.13	2	62651.5	+1	$4p^3\ ^2D^o_{3/2}-73566$	1510.70	6*h*	66194.5		
1594.39	1*r*	62719.9	−5	$4p^3\ ^2D^o_{5/2}-73635$	1510.41	6*h*	66207.2	+1	$4p^3\ ^2D^o_{3/2}-77121$
1593.60	100*R*	62751.0	−1	$4p^3\ ^4S^o_{3/2}-4d\ ^4P_{5/2}$	1509.70	3	66238.3		
1587.97	20	62973.5	0	$4p^3\ ^2D^o_{5/2}-73566$	1509.60	3	66242.7	−4	$4p^3\ ^2D^o_{5/2}-76835$
1583.90	2	63135.3	+2	$4p^3\ ^2D^o_{5/2}-74049$	1506.99	2	66357.4	−5	$4p^3\ ^2D^o_{3/2}-77272$
1575.87	20	63457.0	−2	$4p^3\ ^2D^o_{3/2}-74049$	1504.09	2	66485.4	+2	$4p^3\ ^4S^o_{3/2}-4d\ ^2P_{1/2}$
1574.72	30*r*	63503.4	−2	$4p^3\ ^4S^o_{3/2}-4d\ ^4P_{1/2}$	1503.47	1*r*,*h*	66512.8		
1573.85	60	63538.5	0	$4p^3\ ^2D^o_{3/2}-4d''\ ^2D_{3/2}$	1492.81	1*r*	66987.8		
1572.19	1*r*,*h*	63605.5			1492.34	4	67008.9	+2	$4p^3\ ^4S^o_{3/2}-6s\ ^4P_{5/2}$
1566.39	15	63841.1			1484.36	5*h*	67369.1		
1565.05	7	63895.7	−2	$4p^3\ ^2D^o_{3/2}-4d''\ ^2D_{5/2}$	1472.32	4	67920.0	−2	$4p^3\ ^4S^o_{3/2}-6s\ ^4P_{3/2}$
1562.95	10	63981.6	−1	$4p^3\ ^4S^o_{3/2}-4d\ ^4P_{3/2}$	1470.09	3*r*	68023.0		
1559.53	4	64121.9	−1	$4p^3\ ^4S^o_{3/2}-4d\ ^2P_{1/2}$	1468.42	2*r*	68100.4		
1559.48	4	64124.1			1464.10	2*r*	68301.1	+5	$4p^3\ ^4S^o_{3/2}-4p^4\ ^2D_{1/2}$
1558.28	4*d*+Cu?	64173.3			1461.92	3*r*,*g*?	68403.0	+3	$4p^3\ ^4S^o_{3/2}-4p^4\ ^2D_{3/2}$
1557.20	30	64217.8	−2	$4p^3\ ^2D^o_{5/2}-4d''\ ^2D_{5/2}$	1442.69	5	69315.0	+2	$4p^3\ ^4S^o_{3/2}-6s\ ^4P_{5/2}$
1556.14	40	64261.6	−2	$4p^3\ ^2D^o_{5/2}-75176$	1434.77	3*r*	69697.6	−7	$4p^3\ ^4S^o_{3/2}-69698$
1554.19	4	64342.2	−2	$4p^3\ ^4S^o_{3/2}-4d\ ^2D_{3/2}$	1434.66	1*r*,*h*	69702.9		
1547.59	3	64616.6			1422.95	2*r*,*h*	70276.5		
1546.45	2	64664.2	+1	$4p^3\ ^2D^o_{3/2}-75578$	1412.95	2*r*,*h*	70773.9		
1542.94	4	64811.3	−2	$4p^3\ ^4S^o_{3/2}-4p^4\ ^2P_{1/2}$	1407.34	2*r*	71056.0		
1538.79	20	64986.1	−1	$4p^3\ ^2D^o_{1/2}-75578$					
1534.65	4	65161.4	−2	$4p^3\ ^2D^o_{3/2}-76076$					
1533.67	3	65203.1							
1533.60	3	65206.0	−9	$4p^3\ ^2D^o_{1/2}-75799?$					
1526.78	2	65497.3	−1	$4p^3\ ^4S^o_{3/2}-4p^4\ ^2P_{3/2}$					
1521.81	1*r*	65711.2							
1518.10	1*r*	65871.8							

III. Analysis of the First Spectrum of Arsenic

The above-described efforts to obtain a better description of the As I spectrum were made primarily for the purpose of establishing atomic energy levels characteristic of neutral arsenic atoms and to determine the spectroscopic ionization potential.

1. Observed As I Lines

The first two columns of table 1 contain the observational material upon which a new analysis is based, but as atomic energy levels are directly proportional to frequencies or wave numbers, the wavelengths in column 1 were converted to vacuum wave numbers in column 3. Wave lengths shorter than 2000 A, being vacuum values, were converted to vacuum wave numbers by calculating the reciprocals [14] of the observed wavelengths. Vacuum wave numbers for lines between 2000 and 10000 A were taken directly from Kayser's Table of Wave Numbers [15], but for lines of greater wavelength the values were obtained by computing the reciprocals of the wavelengths after correcting them to vacuum with the aid of the atmospheric dispersion formula of Meggers and Peters [16].

The first regularities among arsenic lines were discovered in 1894 by Kayser and Runge [17], who found 20 ultraviolet lines to be connected by wave-number differences of 461 and 8058. Thirty years later, another significant wave-number difference was found by Ruark, et al. [3], and in 1927 McLennan and McLay [18] interpreted these constant differences as intervals between five low-energy odd levels, namely $^4S^o_{3/2}$, $^2D^o_{3/2}$, $^2D^o_{5/2}$, $^2P^o_{3/2}$, $^2P^o_{1/2}$, arising from the electron configuration $4s^2\,4p^3$. Naturally, the separations of these five levels were exploited in each subsequent attempt to extend the analysis of the As I spectrum; they have accounted for most of the observed ultraviolet lines and coincidentally disclosed a considerable number of high-energy even levels, the lowest of which are important for the classification of visible and infrared lines, thus revealing other high-energy odd levels. In the present instance, the separations of the five lowest odd levels (0.0, 10592.5, 10914.6, 18186.1, 18647.5 cm^{-1}), when applied to the wave numbers of ultraviolet As I lines, yielded about 50 even levels (50693.8 to 77272.5 cm^{-1}), and the separations of seven lowest even levels (50693.8, 51610.2, 52897.9, 53135.6, 54605.3, 60815.0, 60834.8 cm^{-1}), when applied to the wave numbers of visible and infrared lines, led to some 25 high-energy odd levels (60791.5 to 73244.8 cm^{-1}). It is seen that the latter set of odd levels is completely interlaced and surrounded by the even levels with which they combine. Consequently, only fringe combinations of these odd and even levels have been observed, and a considerable number of permitted combinations must exist in the unexplored infrared.

The inner quantum number (J-value) assigned to each level was originally based on the assumption that the first five odd levels have J-values 1½ 1½, 2½, 0½, 1½, respectively. These and the values for four even levels were verified in 1934 by

Zeeman patterns observed by Green and Barrows [8]. Because all As I levels of unlike parity are permitted to combine if their J-values differ by 0 or by ± 1, the J-values of new levels are usually fixed unambiguously by the observed combinations, but there are still a few high levels whose J-values are uncertain by one unit.

After establishing a large number of atomic energy levels, mostly with definite J-values, the next ambition was to group and interpret these as spectral terms characteristic of various configurations of the valence electrons. Aside from eight levels identified by Zeeman effect, the remainder could be grouped and named only by applying interval and intensity rules and by comparison with analogous spectra of other elements, in particular with the N I [19], P I [19], and Se II [20] spectra, which were not well known 20 years ago. This procedure consumed all the established As I levels except a few very high ones, which are retained as "miscellaneous levels" because they may represent fragments of terms approaching the ionization limit. At this stage it was possible to discard some false levels involving doubly classified lines, and also to add some levels resting on a single combination, as for example those with J-value 3½. For fixing the latter, strong lines, otherwise unclassified, were usually found in their expected places.

2. As I Multiplets

As a check on the consistency and plausibility of spectral term designations, the square arrays of consecutive level combinations were rewritten to display As I multiplets and electron configurations so that level separations and line intensities could be easily seen. These combinations, grouped into multiplets, are presented in table 2 (ultraviolet) and table 3 (infrared and visible). All of the observed combinations are included in these tables except three that account for five lines due to transitions from $5p\ ^4D^\circ$ to higher levels. From tables 2 and 3 stem the "Term combination" entered in column 5 of table 1 for each classified As I line, and also the difference between observed and calculated wave number shown in column 4 of table 1.

TABLE 2. As I *combination of odd and even terms*

Desig.	Desig. Level	$4p^3\ ^4S^\circ_{1\frac{1}{2}}$ *0.0*	$4p^3\ ^2D^\circ_{2\frac{1}{2}}$ *10914.6*	$4p^3\ ^2D^\circ_{1\frac{1}{2}}$ *10592.5*	$4p^3\ ^2P^\circ_{1\frac{1}{2}}$ *18647.5*	$4p^3\ ^2P^\circ_{0\frac{1}{2}}$ *18186.1*
$4p^4\ ^4P_{2\frac{1}{2}}$	55366.4	55366.4 (200)				
$4p^4\ ^4P_{1\frac{1}{2}}$	56863.6	56863.6 (100)				
$4p^4\ ^4P_{0\frac{1}{2}}$	57488.1	*57488.1 (60)				
$4p^4\ ^2D_{2\frac{1}{2}}$	68402.7	68403.0 (3*eq*?)	*57488.1 (60)	57810.2 (30)	49755.1 (100*r*)	
$4p^4\ ^2D_{1\frac{1}{2}}$	68300.6	68301.1 (2*r*)	57385.8 (10)	57708.1 (30)	49653.1 (100)	50114.5 (100*r*)
$4p^4\ ^2P_{1\frac{1}{2}}$	64811.5	64811.3 (4)	53897.0 (10)	54219.3 (40)	46163.8 (150)	46625.3 (100)
$4p^4\ ^2P_{0\frac{1}{2}}$	65497.4	65497.3 (2)		54905.2 (2)	46849.9 (50)	47311.3 (100)
$4p^4\ ^2S_{0\frac{1}{2}}$	66588.3			55996.1 (3)	47940.6 (30)	48402.2 (50)
$5s\ ^4P_{2\frac{1}{2}}$	52897.9	52898.2 (2000*R*)	41983.2 (150*r*)	42305.2 (10)	34250.4 (2)	
$5s\ ^4P_{1\frac{1}{2}}$	51610.2	51610.4 (1500*R*)	40695.5 (200*r*)	41017.7 (50)	32962.7 (40)	33424.0 (20)
$5s\ ^4P_{0\frac{1}{2}}$	50693.8	50694.0 (1000*R*)		40101.7 (50)	32046.1 (50)	32507.5 (20)
$5s\ ^2P_{1\frac{1}{2}}$	54605.3	54606.0 (50)	43690.5 (500*R*)	44012.8 (50)	35957.8 (200*r*)	36419.1 (50*r*)
$5s\ ^2P_{0\frac{1}{2}}$	53135.6	53136.1 (40)		42543.0 (500*R*)	34488.0 (50*r*)	34949.4 (100*r*)
$5s'\ ^2D_{2\frac{1}{2}}$	60815.0	60815.0 (5)	49900.4 (300*r*)	50222.7 (100*r*)	42167.5 (100*r*)	
$5s'\ ^2D_{1\frac{1}{2}}$	60834.8	60835.0 (2)	49920.3 (20)	50242.4 (200*r*)	42187.1 (80*r*)	42648.5 (50)
$5s''\ ^2S_{0\frac{1}{2}}$	72399.4				53751.9 (80)	54213.2 (40)
$6s\ ^4P_{2\frac{1}{2}}$	69314.8	69315.0 (5)	58400.3 (10)	58722.0 (3)	50667.5 (3)	
$6s\ ^4P_{1\frac{1}{2}}$	67920.2	67920.0 (4)	57005.7 (10)		49272.8 (2)	49734.0 (20)
$6s\ ^4P_{0\frac{1}{2}}$	67008.7	67008.9 (4)		56416.2 (8)	48361.2 (20)	48822.6 (50)
$6s\ ^2P_{1\frac{1}{2}}$	69696.2		58781.3 (30)		51048.8 (40*r*)	51510.3 (5)
$6s\ ^2P_{0\frac{1}{2}}$	68314.5			57722.1 (30)	49666.9 (15)	50128.3 (20)
$4d\ ^4P_{2\frac{1}{2}}$	62751.1	62751.0 (100*R*)	51836.6 (3)	52159.1 (20)	44103.3 (25)	
$4d\ ^4P_{1\frac{1}{2}}$	63503.6	63503.4 (30*r*)	52589.0 (5)	52911.4 (5)	44856.0 (20)	45317.4 (15)
$4d\ ^4P_{0\frac{1}{2}}$	63981.7	63981.6 (10)		53389.6 (40)	45334.0 (10)	45795.5 (20)
$4d\ ^2F_{3\frac{1}{2}}$	66785.2		55870.6 (50)			
$4d\ ^2F_{2\frac{1}{2}}$	64639.5			54047.0 (40)		
$4d\ ^2D_{2\frac{1}{2}}$	66725.4		55810.7 (40)	56133.1 (50)	48077.7 (7)	
$4d\ ^2D_{1\frac{1}{2}}$	64342.4	64342.2 (4)	53427.9 (30)	53750.1 (80)	45694.7 (5)	
$4d\ ^2P_{1\frac{1}{2}}$	66485.2	66485.4 (2)	55570.7 (2)	55892.8 (15)	47837.6 (6)	48298.8 (30)
$4d\ ^2P_{0\frac{1}{2}}$	64122.0	64121.9 (4)			45474.6 (5)	45935.9 (5)
$4d'\ ^2P_{1\frac{1}{2}}$	72125.0		61210.4 (5)	61532.0 (1)	53477.6 (15)	53938.9 (3)
$4d'\ ^2P_{0\frac{1}{2}}$	71924.4			61331.6 (4)	53277.1 (2)	53738.3 (5)
$4d''\ ^2D_{2\frac{1}{2}}$	74453.1		63538.5 (60)			
$4d''\ ^2D_{1\frac{1}{2}}$	74810.5		63895.7 (7)	64217.8 (30)	56163.4 (50)	

TABLE 2. As I *combination of odd and even terms*—Continued

Desig.	Desig. Level	$4p^5\ {}^4S_{1\frac{1}{2}}$ 0.0	$4p^5\ {}^2D_{2\frac{1}{2}}$ 10014.6	$4p^5\ {}^2D_{1\frac{1}{2}}$ 10592.5	$4p^5\ {}^2P_{1\frac{1}{2}}$ 18647.5	$4p^5\ {}^2P_{0\frac{1}{2}}$ 18786.1
	68644.8				50997.2 (2)	51458.9 (2c)
	68688.3	68697.6 (3c)	58783.4 (15)	59106.2 (7)	51051.1 (20)	
	71055.2		60141.0 (2)	60462.4 (4)	52407.3 (1)	52868.9 (5)
	72519.4		61604.4 (2)	61926.5 (5)	53872.3 (2)	54333.6 (10)
	72779.0		61864.5 (6)	62186.3 (3)	54132.5 (10)	54592.9 (30)
	73566.0		62651.5 (2)	62973.5 (20)		
	73635.0		62719.9 (1e)		54987.7 (5h)	55449.0 (10h)
	74049.7		63135.3 (2)	63457.0 (20)		
	74143.8				55496.4 (3)	55957.5 (6)
	75086.0				56438.5 (2)	56900.0 (4)
	75176.4		64261.6 (40)		†56530.1? (5)	
	75578.7		64664.2 (2)	64986.1 (20)	56931.1 (2h)	
	75799.4			65206.0 (3)	57152.0 (1eh)	57613.6 (3)
	76076.2		65161.4 (4)		57429.0 (10)	57891.1 (4)?
*	76835.6		65920.9 (1e)	66242.7 (3)	58188.6 (1)	
	77121.7		66207.2 (6h)		58474.1 (7h)	
	77272.5		66357.4 (2)		58625.6 (4)	

*Blend.
†Blend with As II?.

TABLE 3. As I *Combinations of even and odd terms*

Desig	Desig Level	$5s\ {}^4P_{1\frac{1}{2}}$ 52897.9	$5s\ {}^4P_{2\frac{1}{2}}$ 51610.2	$5s\ {}^4P_{0\frac{1}{2}}$ 50693.8	$5s\ {}^2P_{1\frac{1}{2}}$ 54605.3	$5s\ {}^2P_{0\frac{1}{2}}$ 53135.6	$5s'\ {}^2D_{2\frac{1}{2}}$ 60815.0	$5s'\ {}^2D_{1\frac{1}{2}}$ 60834.8
$5p\ {}^4D_{3\frac{1}{2}}$	62871.2	9973.30 (400)						
$5p\ {}^4D_{2\frac{1}{2}}$	61885.0	8787.19 (4)	10074.78 (400)					
$5p\ {}^4D_{1\frac{1}{2}}$	60860.0		9249.78 (50)	10166.26 (300)				
$5p\ {}^4D_{0\frac{1}{2}}$	60791.5		9181.22 (20)	10097.68 (150)				
$5p\ {}^4P_{2\frac{1}{2}}$	63282.8	10384.93 (400)	11672.62 (100)		8677.54 (10)			
$5p\ {}^4P_{1\frac{1}{2}}$	62398.0		10787.70 (150)	11704.20 (50)		9262.42 (8)		
$5p\ {}^4P_{0\frac{1}{2}}$	62026.3		10416.04 (300)	11332.41 (150)		8890.79 (10)		
$5p\ {}^4S_{1\frac{1}{2}}$	63647.0	10749.04 (250)	12036.75 (50)			10511.4 (3)		
$5p\ {}^2D_{2\frac{1}{2}}$	64169.2	11271.30 (100)	12558.94 (30)		9563.93 (100)			
$5p\ {}^2D_{1\frac{1}{2}}$	62554.4		10944.17 (50)	11860.66 (100)		9418.87 (200)		
$5p\ {}^2P_{1\frac{1}{2}}$	64687.6				10082.24 (200)	11551.99 (100)		
$5p\ {}^2P_{0\frac{1}{2}}$	64323.8		12713.61 (4)		9718.49 (50)	11188.17 (50)		
$5p\ {}^2S_{0\frac{1}{2}}$	64059.0				9453.65 (100)			
$5p'\ {}^2F_{3\frac{1}{2}}$	70927.5						10112.49 (100)	
$5p'\ {}^2F_{2\frac{1}{2}}$	70821.4		19211.24 (4)				10006.40 (5)	9986.65 (100)
$5p'\ {}^2D_{2\frac{1}{2}}$	70876.6				15771.16 (40)		9561.60 (20)	9541.83 (4)
$5p'\ {}^2D_{1\frac{1}{2}}$	70108.4							9273.59 (5)
$5p'\ {}^2P_{1\frac{1}{2}}$	71931.7		20321.65 (2)			18796.16 (5)	11116.65 (20)	11097.02 (2)
$5p'\ {}^2P_{0\frac{1}{2}}$	71796.2		20185.96 (10)		17190.97 (2h)	18660.53 (1h?)		10961.52 (10)
$6p\ {}^4D_{3\frac{1}{2}}$	72341.6	19443.68 (100)						
$6p\ {}^4D_{2\frac{1}{2}}$	71131.0		19520.72 (100)		16525.78 (5h)		10316.20 (10)	
$6p\ {}^4D_{1\frac{1}{2}}$	70516.6			19822.74 (50)		17380.90 (5h)	9701.54 (5)	9681.88 (10)
$6p\ {}^4D_{0\frac{1}{2}}$	70974.9			19581.09 (30) Cu*				
$6p\ {}^4P_{2\frac{1}{2}}$	72501.9	19603.97 (60)	20891.64 (3k)					
$6p\ {}^4P_{1\frac{1}{2}}$	70849.7		19239.49 (50)					
$6p\ {}^4P_{0\frac{1}{2}}$	69881.5			19187.68 (20)				
$6p\ {}^4S_{1\frac{1}{2}}$	72680.3	19722.35 (100)	21010.14 (5k)			19484.80 (80)		
$6p\ {}^2D_{2\frac{1}{2}}$	72944.4	20046.48 (6h)			18339.10 (150)		12129.43 (20)	
$6p\ {}^2D_{1\frac{1}{2}}$	71621.4		20011.13 (8)	20927.72 (1h)		18485.56 (100)		(10786.6)§
$6p\ {}^2P_{1\frac{1}{2}}$	73244.8				18639.23 (60)		12429.83 (8)	12410.13 (5h)
$6p\ {}^2P_{0\frac{1}{2}}$	72800.6				18195.31 (5)	19665.02 (10)		11965.82 (10)
$6p\ {}^2S_{0\frac{1}{2}}$	71189.1		19588.84 (20)			18063.57 (3)		10364.4 (3)

§ Calculated wave numbers line masked.

TABLE 4.—As I *predicted and observed terms*

Electron configuration	Limit As II	Predicted	Observed
$4s^2\ 4p^3$		$^4S^\circ\ ^2(P^\circ\ D^\circ)$	$4p^3\ ^4S^\circ$, $4p^3\ ^2P^\circ$, $4p^3\ ^2D^\circ$
$4s\ \ 4p^4$		$^4P\ ^2(S\ P\ D)$	$4p^4\ ^4P$, $4p^4\ ^2S$, $4p^4\ ^2P$, $4p^4\ ^2D$
$4p^5$		$^2P^\circ$	
$4s^2\ 4p^2\ ns$	^3P	$^4P\ ^2P$	$5s\ ^4P$; $6s\ ^4P$, $5s\ ^2P$, $6s\ ^2P$
	^1D	^2D	$5s'\ ^2D$
	^1S	^2S	$5s''\ ^2S$
$4s^2\ 4p^2\ np$	^3P	$^4(S^\circ\ P^\circ\ D^\circ)\ ^2(S^\circ\ P^\circ\ D^\circ)$	$\{5p\ ^4S^\circ,\ 6p\ ^4S^\circ,\ 5p\ ^4P^\circ,\ 6p\ ^4P^\circ,\ 5p\ ^4D^\circ,\ 6p\ ^4D^\circ$ $\{5p\ ^2S^\circ,\ 6p\ ^2S^\circ,\ 5p\ ^2P^\circ,\ 6p\ ^2P^\circ,\ 5p\ ^2D^\circ,\ 6p\ ^2D^\circ$
	^1D	$^2(P^\circ\ D^\circ\ F^\circ)$	$5p'\ ^2P^\circ,\ 5p'\ ^2D^\circ,\ 5p'\ ^2F^\circ$
	^1S	$^2P^\circ$	
$4s^2\ 4p^2\ nd$	^3P	$^4(P\ D\ F)\ ^2(P\ D\ F)$	$4d\ ^4P$, $4d\ ^2P$, $4d\ ^2D$, $4d\ ^2F$
	^1D	$^2(S\ P\ D\ F\ G)$	$4d'\ ^2P$
	^1S	^2D	$4d''\ ^2D$

3. As I Spectral Terms

Arsenic has atomic number 33, and the neutral atoms have electronic structure as follows: $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^3$. The first spectrum of arsenic arises from changes in atomic energy resulting from alterations of the principal quantum numbers and total momenta of the last-named five valence electrons. The theoretical spectral terms associated with various configurations of electrons [21] are listed in table 4.

It was stated above that the empirically established atomic energy levels derived from analysis of the As I spectrum occur in three groups (1) five odd levels of low energy content (including zero, the normal state), (2) about 50 even levels of higher energy, and (3) some 25 odd levels intermingled with group (2). That group 1 is identified with the normal configuration $4s^2\ 4p^3$ is verified by many experimental facts including line intensities, self-reversals, and Zeeman patterns. Group 2 arises in part from the configuration $4s\ 4p^4$, which yields eight levels, $^4P_{2\frac{1}{2}}$, $^4P_{1\frac{1}{2}}$, $^4P_{0\frac{1}{2}}$, $^2S_{0\frac{1}{2}}$, $^2P_{0\frac{1}{2}}$, $^2P_{1\frac{1}{2}}$, $^2D_{1\frac{1}{2}}$, $^2D_{2\frac{1}{2}}$, all of which have been selected with the aid of analogous spectra as guides. The remainder of group 2 must be ascribed to $4s^2\ 4p^2\ ns$ and $4s^2\ 4p^2\ nd$, whereas group 3 can originate only in $4s^2\ 4p^2\ np$. There is, naturally, no difficulty about distinguishing even and odd levels, but the proper choice of even levels for terms due to ns and nd electrons added to $4s^2\ 4p^2$ is complicated by overlapping of energies of the different electron configurations. These configurations give rise respectively to 8 and 28 levels, which converge toward five limits ($^3P_0$, $^3P_1$, $^3P_2$, $^1D_2$, $^1S_0$) in the As II spectrum when the running electron is removed by ionization. According to Rao [13], these levels in the As II spectrum have the following values in cm^{-1}: $^3P_0=0$; $^3P_1=1063$, $^3P_2=2540$, $^1D_2=10095$, $^1S_0=22599$. Guided by these intervals in convergence limits, we have provisionally interpreted the observed As I even levels as shown in table 4. Although the configuration $4s^2\ 4p^3$ yields only five levels, including the ground term of As I, if the total quantum number is altered for one of the p electrons so that they are not all equivalent, the configuration produces no less than 21 levels, which converge also to the above mentioned five limits in As II. Terms associated with $5p$ and $6p$ electrons have been identified, as indicated in table 4. Ten hazy lines from 4910.23 to 4301.56 A could not be classified; they may represent some leading combinations of terms arising from $7p$ electrons with those from $5s$.

Full particulars (except g values) regarding the now known terms of the As I spectrum are presented in table 5, where columns 1 to 5 contain (1) electron configuration (and convergence limit), (2) term designation, (3) level value (relative to $4p^3\ ^4S^\circ_{1\frac{1}{2}}=0.0$, (4) intervals between levels of terms, and (5) combinations observed for each term.

4. First Ionization Potential of Arsenic

In 1942, before the ultraviolet portion of the As I spectrum had been satisfactorily observed, an attempt was made to obtain a value of the effective quantum number of the $5s$ electron for As by comparing the values of neighboring elements and thus arrive at an estimate of the ionization potential. The result was not published, but it will be given here because it adds credence to the value recently derived from spectral series. The effective quantum numbers, $n^*=\sqrt{R/T}$, corresponding to $5s$ in spectra of elements flanking As are shown in table 6.

By interpolation, the $5s$ value of n^* for As I is found to be about 1.960. The corresponding term, $T=R/n^{*2}$, has a numerical value of $109737/(1.960)^2=28565$, and if this is added to the wave number 50694 of the resonance line (1972.62 A), a value of 79259

TABLE 5. *Terms of the* As I *spectrum*

Electron configuration	Term designation	Level cm⁻¹	Interval cm⁻¹	Combinations
1	2	3	4	5
$4s^2\ 4p^3$	$4p^3\ {}^4S_{1\frac{1}{2}}$	*0.0*		$4p^4\ {}^3P$, $5s\ {}^4P$, $6s\ {}^4P$, $4d\ {}^4P$, $4p^4\ {}^3P$, $5s\ {}^3P$, $4d\ {}^3P$, $4p^4\ {}^3D$, $5s'\ {}^3D$, $4d\ {}^3D$, 69698.
$4s^2\ 4p^3$	$4p^3\ {}^2D_{1\frac{1}{2}}$	*10592.5*	322.1	$5s\ {}^4P$, $6s\ {}^4P$, $4d\ {}^4P$, $4p^4\ {}^2S$, $4p^4\ {}^2P$, $5s\ {}^2P$, $6s\ {}^2P$, $4d\ {}^2P$, $4d'\ {}^2P$, $4p^4\ {}^2D$, $5s'\ {}^2D$, $4d\ {}^2D$, $4d''\ {}^2D$, $4d\ {}^2F$, 69698, 71055, 72519, 72779, 73566, 73635, 74049, 75176, 75578, 75799, 76076, 76835, 77121, 77272.
	$4p^3\ {}^2D_{2\frac{1}{2}}$	*10914.6*		
$4s^2\ 4p^3$	$4p^3\ {}^2P_{0\frac{1}{2}}$	*18186.1*	461.4	$5s\ {}^4P$, $6s\ {}^4P$, $4d\ {}^4P$, $4p^4\ {}^2S$, $5s''\ {}^2S$, $4p^4\ {}^2P$, $5s\ {}^2P$, $6s\ {}^2P$, $4d\ {}^2P$, $4d'\ {}^2P$, $4p^4\ {}^2D$, $5s'\ {}^2D$, $4d\ {}^2D$, $4d''\ {}^2D$, 69644, 69698, 71055, 72519, 72779, 73635, 74143, 75086, 75176, 75578, 75799, 76076, 76835, 77121, 77272.
	$4p^3\ {}^2P_{1\frac{1}{2}}$	*18647.5*		
$4s^2\ 4p^2(^3P)5s$	$5s\ {}^4P_{0\frac{1}{2}}$	50693.8	916.4	$4p^3\ {}^4S°$, $5p\ {}^4S°$, $6p\ {}^4S°$, $5p\ {}^4P°$, $6p\ {}^4P°$, $5p\ {}^4D°$, $6p\ {}^4D°$, $6p\ {}^2S°$, $4p^3\ {}^2P°$, $5p\ {}^2P°$, $5p'\ {}^2P°$, $4p^3\ {}^2D°$, $5p\ {}^2D°$, $6p\ {}^2D°$, $5p'\ {}^2F°$.
	$5s\ {}^4P_{1\frac{1}{2}}$	51610.2	1287.7	
	$5s\ {}^4P_{2\frac{1}{2}}$	52897.9		
$4s^2\ 4p^2(^3P)5s$	$5s\ {}^2P_{0\frac{1}{2}}$	53135.6	1469.7	$4p^3\ {}^4S°$, $5p\ {}^4S°$, $6p\ {}^4S°$, $5p\ {}^4P°$, $6p\ {}^4D°$, $5p\ {}^2S°$, $6p\ {}^2S°$, $4p^3\ {}^2P°$, $5p\ {}^2P°$, $6p\ {}^2P°$, $5p'\ {}^2P°$, $4p^3\ {}^2D°$, $5p\ {}^2D°$, $6p\ {}^2D°$, $5p'\ {}^2D°$.
	$5s\ {}^2P_{1\frac{1}{2}}$	54605.3		
$4s\ \ 4p^4$	$4p^4\ {}^4P_{2\frac{1}{2}}$	55366.4	−1497.2	$4p^3\ {}^4S°$.
	$4p^4\ {}^4P_{1\frac{1}{2}}$	56863.6	−624.5	
	$4p^4\ {}^4P_{0\frac{1}{2}}$	57488.1		
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^4D_{3\frac{1}{2}}$	*60791.5*	68.5	$5s\ {}^4P$, $4d'\ {}^2P$, $4d''\ {}^2D$, 75176.
	$5p\ {}^4D_{2\frac{1}{2}}$	*60860.0*	825.0	
	$5p\ {}^4D_{1\frac{1}{2}}$	*61685.0*	1186.2	
	$5p\ {}^4D_{0\frac{1}{2}}$	*62871.2*		
$4s^2\ 4p^2(^1S)5s$	$5s'\ {}^2D_{2\frac{1}{2}}$	60815.0	−19.8	$4p^3\ {}^4S°$, $6p\ {}^4D°$, $6p\ {}^2S°$, $4p^3\ {}^2P°$, $6p\ {}^2P°$, $5p'\ {}^2P°$, $4p^3\ {}^2D°$, $6p\ {}^2D°$, $5p'\ {}^2D°$, $5p'\ {}^2F°$.
	$5s'\ {}^2D_{1\frac{1}{2}}$	60834.8		
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^4P_{0\frac{1}{2}}$	*62026.3*	371.7	$5s\ {}^4P$, $5s\ {}^2P$.
	$5p\ {}^4P_{1\frac{1}{2}}$	*62398.0*	884.8	
	$5p\ {}^4P_{2\frac{1}{2}}$	*63282.8*		
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^2D_{1\frac{1}{2}}$	*62554.4*	1614.8	$5s\ {}^4P$, $5s\ {}^2P$, $4d'\ {}^2P$.
	$5p\ {}^2D_{2\frac{1}{2}}$	*64169.2*		
$4s^2\ 4p^2(^3P)4d$	$4d\ {}^4P_{2\frac{1}{2}}$	62751.1	−732.5	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$4d\ {}^4P_{1\frac{1}{2}}$	63503.6	−478.1	
	$4d\ {}^4P_{0\frac{1}{2}}$	63981.7		
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^2S_{0\frac{1}{2}}$	*63647.0*		$5s\ {}^4P$, $5s\ {}^2P$.
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^2S_{0\frac{1}{2}}$	*64059.0*		$5s\ {}^2P$.
$4s^2\ 4p^2(^3P)4d$	$4d\ {}^2P_{0\frac{1}{2}}$	64122.0	2363.2	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$4d\ {}^2P_{1\frac{1}{2}}$	66485.2		
$4s^2\ 4p^2(^3P)5p$	$5p\ {}^2P_{0\frac{1}{2}}$	*64323.8*	363.8	$5s\ {}^4P$, $5s\ {}^2P$.
	$5p\ {}^2P_{1\frac{1}{2}}$	*64687.6*		
$4s^2\ 4p^2(^3P)4d$	$4d\ {}^2D_{1\frac{1}{2}}$	64342.4	2383.0	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$4d\ {}^2D_{2\frac{1}{2}}$	66725.4		
$4s^2\ 4p^2(^3P)4d$	$4d\ {}^2F_{3\frac{1}{2}}$	64639.5	2145.7	$4p^3\ {}^2D°$.
	$4d\ {}^2F_{3\frac{1}{2}}$	66785.2		
$4s\ \ 4p^4$	$4p^4\ {}^2P_{1\frac{1}{2}}$	64811.5	−685.9	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$4p^4\ {}^2P_{0\frac{1}{2}}$	65497.4		
$4s\ \ 4p^4$	$4p^4\ {}^2S_{0\frac{1}{2}}$	66588.3		$4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
$4s^2\ 4p^2(^3P)6s$	$6s\ {}^4P_{0\frac{1}{2}}$	67008.7	911.5	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$6s\ {}^4P_{1\frac{1}{2}}$	67920.2	1394.6	
	$6s\ {}^4P_{2\frac{1}{2}}$	69314.8		
$4s\ \ 4p^4$	$4p^4\ {}^2D_{1\frac{1}{2}}$	68300.6	102.1	$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$4p^4\ {}^2D_{2\frac{1}{2}}$	68402.7		
$4s^2\ 4p^2(^3P)6s$	$6s\ {}^2P_{0\frac{1}{2}}$	68314.5	1381.7	$4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
	$6s\ {}^2P_{1\frac{1}{2}}$	69696.2		
	$0\frac{1}{2}, 1\frac{1}{2}$	69644.8		$4p^3\ {}^2P°$.
	$1\frac{1}{2}, 2\frac{1}{2}$	69698.3		$4p^3\ {}^4S°$, $4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
$4s^2\ 4p^2(^3P)6p$	$6p\ {}^4P_{0\frac{1}{2}}$	*69881.5*	968.2	$5s\ {}^4P$.
	$6p\ {}^4P_{1\frac{1}{2}}$	*70849.7*	1652.2	
	$6p\ {}^4P_{2\frac{1}{2}}$	*72501.9*		
$4s^2\ 4p^2(^1D)5p$	$5p'\ {}^2D_{1\frac{1}{2}}$	*70108.4*	268.2	$5s\ {}^2P$, $5s'\ {}^2D$.
	$5p'\ {}^2D_{2\frac{1}{2}}$	*70376.6*		
$4s^2\ 4p^2(^3P)6p$	$6p\ {}^4D_{3\frac{1}{2}}$	*70874.9*	241.7	$5s\ {}^4P$, $5s\ {}^2P$, $5s'\ {}^2D$.
	$6p\ {}^4D_{2\frac{1}{2}}$	*70516.6*	614.4	
	$6p\ {}^4D_{1\frac{1}{2}}$	*71131.0*	1210.6	
	$6p\ {}^4D_{0\frac{1}{2}}$	*72341.6*		
$4s^2\ 4p^2(^1D)5p$	$5p'\ {}^2F_{3\frac{1}{2}}$	*70821.4*	106.1	$5s\ {}^4P$, $5s'\ {}^2D$.
	$5p'\ {}^2F_{2\frac{1}{2}}$	*70927.5*		
	$1\frac{1}{2}$	71055.2		$4p^3\ {}^2P°$, $4p^3\ {}^2D°$.
$4s^2\ 4p^2(^3P)6p$	$6p\ {}^2S_{0\frac{1}{2}}$	*71190.1*		$5s\ {}^4P$, $5s\ {}^2P$, $5s'\ {}^2D$.

TABLE 5. *Terms of the* As I *spectrum*—Continued

Electron configuration	Term designation	Level cm^{-1}	Interval cm^{-1}	Combinations
1	2	3	4	5
$4s^2\ 4p^3(^3P)6p$	$6p\ ^3D_{1\frac{1}{2}}$	*71681.4*	1323.0	$5s\ ^4P,\ 5s\ ^2P,\ 5s'\ ^2D.$
	$6p\ ^3D_{2\frac{1}{2}}$	*72944.4*		
$4s^2\ 4p^3(^1D)5p$	$5p'\ ^2P^\circ_{1\frac{1}{2}}$	*71796.2*	135.5	$5s\ ^4P,\ 5s\ ^2P,\ 5s'\ ^2D.$
	$5p'\ ^2P^\circ_{2\frac{1}{2}}$	*71931.7*		
$4s^2\ 4p^3(^1D)4d$	$4d'\ ^2P_{1\frac{1}{2}}$	71924.4	200.6	$5p\ ^4D^\circ,\ 4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ,\ 5p\ ^2D^\circ.$
	$4d'\ ^2P_{2\frac{1}{2}}$	72125.0		
$4s^2\ 4p^3(^1S)5s$	$5s''\ ^2S_{0\frac{1}{2}}$	72399.4		$4p^3\ ^2P^\circ.$
	$1\frac{1}{2}$	72519.4		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
$4s^2\ 4p^3(^3P)6p$	$6p\ ^4S^\circ_{1\frac{1}{2}}$	*72680.5*		$5s\ ^4P,\ 5s\ ^2P.$
	$1\frac{1}{2}$	72779.0		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
$4s^2\ 4p^3(^3P)6p$	$6p\ ^2P^\circ_{0\frac{1}{2}}$	*72800.6*	444.2	$5s\ ^2P,\ 5s'\ ^2D.$
	$6p\ ^2P^\circ_{1\frac{1}{2}}$	*73244.8*		
	$1\frac{1}{2},\ 2\frac{1}{2}$	73566.0		$4p^3\ ^2D^\circ.$
	$1\frac{1}{2}$	73635.0		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	74049.7		$4p^3\ ^2D^\circ.$
	$0\frac{1}{2},\ 1\frac{1}{2}$	74143.8		$4p^3\ ^2P^\circ.$
$4s^2\ 4p^3(^1S)4d$	$4d''\ ^2D_{2\frac{1}{2}}$	74453.1	−357.4	$5p\ ^4D^\circ,\ 4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$4d''\ ^2D_{1\frac{1}{2}}$	74810.5		
	$0\frac{1}{2},\ 1\frac{1}{2}$	75086.0		$4p^3\ ^2P^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	75176.4		$5p\ ^4D^\circ,\ 4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	75578.7		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$0\frac{1}{2},\ 1\frac{1}{2}$	75799.4		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2}$	76076.2		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	76835.6		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	77121.7		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
	$1\frac{1}{2},\ 2\frac{1}{2}$	77272.5		$4p^3\ ^2P^\circ,\ 4p^3\ ^2D^\circ.$
As II($^3P_0$)	Limit	*79165*		

cm^{-1} is obtained for the limit ($^3P_0$ of As II). This corresponds to a value of $79259 \times 1.2395 \times 10^{-4} = 9.82$ ev for the principal ionization potential of arsenic.

In January 1950, after our ultraviolet As lines had been assembled and classified, it was observed that three high even levels (67009, 67920, 69315 cm^{-1}) had the same J-values and nearly the same separations as the three lowest even levels (50694, 51610, 52898 cm^{-1}). Since the latter group had already been identified as $5s\ ^4P_{0\frac{1}{2}, 1\frac{1}{2}, 2\frac{1}{2}}$, it was natural to regard the higher group as $6s\ ^4P_{0\frac{1}{2}, 1\frac{1}{2}, 2\frac{1}{2}}$. While writing this paragraph, we observed that Rao [12] interpreted these levels in exactly this way but nevertheless arrived at an erroneous value of the ionization potential. Because the line intensities corresponding to transitions from these terms to the ground state are in the ratio of several hundred to one, it seems hopeless to detect similar levels of $7s$, and we must be content with two terms of a spectral series. With the aid of a Rydberg interpolation table, the effective quantum numbers and limits were calculated as shown in table 7.

The last column of table 7 contains three values of As I series limits corrected to the first level ($^3P_0$) of As II. The average of these values is 79527 cm^{-1}, which corresponds to 9.85 ev. Experience shows that the Rydberg formula generally yields too large

TABLE 6. *Effective quantum numbers n* for 5s*

Spectrum	Term	T	R/T	$n^* = \sqrt{R/T}$
29 Cu I	$5s\ ^2S_{0\frac{1}{2}}$	19180	5.724	2.3920
30 Zn I	$5s\ ^3S_1$	22094	4.967	2.229
31 Ga I	$5s\ ^2S_{0\frac{1}{2}}$	23592	4.651	2.157
32 Ge I	$5s\ ^3P^\circ_0$	26173	4.193	2.048
33 As I	$5s\ ^4P_{0\frac{1}{2}}$	28565	3.842	1.960
34 Se I	$5s\ ^3S_1$	30476	3.598	1.898
35 Br I	$5s\ ^4P_{2\frac{1}{2}}$	32120	3.416	1.848
36 Kr I	$5s\ ^3P^\circ_2$	32943	3.331	1.825
37 Rb I	$5s\ ^2S_{0\frac{1}{2}}$	33691	3.257	1.805

TABLE 7. As I *effective quantum numbers and series*

Electron	Lines cm^{-1}	Rydberg terms	$n^* = \sqrt{R/T}$	Limits in As II	Lowest limit
$5s$	50694	28948	1.947	$^3P_0 = 79642$	79642
$6s$	67009	12633	2.947		
$5s$	51610	28941	1.947	$^3P_1 = 80551$	79488
$6s$	67920	12631	2.947		
$5s$	52898	29094	1.942	$^3P_2 = 81902$	79452
$6s$	69315	12677	2.942		

a limit, and a more reliable value will result from the use of a Ritz formula of the type,

$$\nu = L - R/(n+\mu+\alpha T)^2.$$

An examination of s-series represented by the Ritz formula in Zn I, Ga I, and Ge I spectra indicates that $\alpha = 2.5 \times 10^{-6}$ for As I. On this assumption, the lowest limits in the final column of table 7 are calculated to be 79185, 79033, and 78995, respectively. The average of these values is 79071 cm^{-1}; it corresponds to 9.80 ev and is in good agreement with 9.82 ev calculated with the interpolated effective quantum number for the 5s electron. We believe that the average of these two figures, 79165 cm^{-1}, corresponding to 9.81 $\pm$0.01 ev, gives a reliable value for the principal ionization potential of arsenic.

IV. References

[1] William F. Meggers and T. L. deBruin, J. Research NBS **3,** 765 (1929) RP116.
[2] L. Bloch and E. Bloch, J. Phys. Rad. [6] **2,** 229 (1921).
[3] A. E. Ruark, F. L. Mohler, P. D. Foote, and R. L. Chenault, BS Sci. Pap. **19,** 463 (1924) S490.
[4] P. Queney, J. Phys. Rad. [6] **10,** 448 (1929).
[5] K. R. Rao, Proc. Roy. Soc. [A] **125,** 238 (1929).
[6] A. S. Rao, Proc. Phys. Soc. London, **44,** 594 (1932).
[7] W. M. Hicks, Phil. Mag. **15,** 1080 (1933).
[8] J. B. Green and W. M. Barrows, Jr., Phys. Rev. **47,** 131 (1935).
[9] H. M. Randall, Astrophys. J. **34,** 1 (1911).
[10] A. G. Shenstone, Phil. Trans. Roy Soc. London [A] **235,** 195, No. 751 (1936); **241,** 297, No. 832 (1948).
[11] J. C. Boyce and H. A. Robinson, J. Opt. Soc. Am. **26,** 133 (1936).
[12] H. Kayser und R. Ritschl, Tabelle der Hauptlinien der Linienspektren aller Elemente. (Julius Springer, Berlin, 1939).
[13] A. S. Rao, Ind. J. Phys. **7,** 561 (1933).
[14] Table of Reciprocals of the Integers from 100,000 through 200,009. (Columbia University Press, New York, N. Y., 1943).
[15] H. Kayser, Tabelle der Schwingungszahlen. (Julius Springer, Berlin 1926).
[16] W. F. Meggers and C. G. Peters, BS Sci. Pap. **14,** 697 (1918) S327.
[17] H. Kayser and C. Runge, Wiedemann's Ann. **52,** 93 (1894).
[18] J. C. McLennan and A. B. McLay, Trans. Roy. Soc. Can. **21,** 3, 63 (1927).
[19] C. E. Moore, Atomic Energy Levels, 1, NBS Circular C467 (1949).
[20] J. C. Van den Bosch, Physica **14,** 249, No. 4 (1948).
[21] F. Hund, Linienspektren und Periodisches System der Elemente. (Julius Springer, Berlin, 1927).

WASHINGTON, May 16, 1950.